

Politikdidaktisches Wissen und Können – eine explorative Studie zur Modellierung und Operationalisierung handlungsnaher Aspekte professioneller Kompetenzen

Von der Fakultät für Gesellschaftswissenschaften
der Universität Duisburg-Essen
zur Erlangung des akademischen Grades

Dr. phil.

genehmigte Dissertation

von

Neumann, Dennis

aus

Essen

1. Gutachterin: Prof. Dr. Manzel, Sabine
2. Gutachterin: Prof. Dr. Oberle, Monika

Tag der Disputation: 16.07.2015

Diese Arbeit wurde von der Fakultät für Gesellschaftswissenschaften der Universität Duisburg-Essen als Dissertation zur Erlangung des Doktorgrades (Dr. phil.) genehmigt.

Name der Gutachterinnen:

1. Prof. Dr. Sabine Manzel
2. Prof. Dr. Monika Oberle

Tag der Disputation: 16.07.2015

Für meine Eltern

Danksagung

Ich danke Prof. Dr. Sabine Manzel für die Betreuung der Promotion und die Begutachtung der Dissertationsschrift. Für die Übernahme des Zweitgutachtens bedanke ich mich bei Prof. Dr. Monika Oberle.

Prof. Dr. Anke Lindmeier gilt mein Dank für die Erlaubnis, das Kompetenzdiagnostik-Programm vKid nutzen zu dürfen sowie für alle konstruktiven Anregungen.

Dorothee Gronostay, Matthias Sowinski, Farina Nagel und Laura Möllers möchte ich für die vielen lehrreichen Diskussionen danken. Ich danke Dorothee Gronostay insbesondere dafür, dass ich die von ihr erhobenen Unterrichtsaufnahmen als Ausgangspunkt für die Erstellung der Videovignetten für die vorliegende Arbeit nutzen durfte. Luisa Kohl und Julia Brüggemann haben als studentische Mitarbeiterinnen wichtige Fleißarbeit geleistet, für die ich mich ebenfalls bedanken möchte.

Herzlich danken möchte ich nicht zuletzt allen Teilnehmerinnen und Teilnehmern der Studie.

Schließlich danke ich dem Zentrum für empirische Bildungsforschung der Universität Duisburg-Essen für die finanzielle Förderung durch Übernahme der Kosten für die Vergütung der Probandinnen und Probanden.

Inhaltsverzeichnis

1 Einleitung	9
2 Politikdidaktisches Wissen und Können von Lehrkräften	16
2.1 Zum professionelles Wissen, Können und Handeln von Lehrkräften	16
2.1.1 Eine Topologie des professionellen Wissens von Lehrkräften	24
2.1.2 Zur mentalen Repräsentation von Wissen	26
2.2 Das politikdidaktische Wissen von Lehrkräften	34
2.2.1 Strukturdimension <i>Facetten</i>	39
2.2.2 Strukturdimension <i>Inhaltsbereiche</i>	52
2.2.3 Strukturdimension <i>Wissensarten</i>	59
2.3 Politikdidaktisches Wissen und Können als Aspekt professioneller Handlungskompetenzen	62
2.3.1 Zum Kompetenzbegriff aus einer kognitionspsychologischen Perspektive ...	62
2.3.1.1 Anwendungsbezug	63
2.3.1.2 Domänenspezifität	65
2.3.1.3 Erlern- und Graduierbarkeit	68
2.3.2 Modellierung von Kompetenzen	68
2.3.3 Messung von Kompetenzen	71
2.3.3.1 Grundlagen und Anforderungen domänenspezifischer Kompetenzdiagnostik	71
2.3.3.2 Möglichkeiten und Grenzen computerbasierter videogestützter Kompetenzdiagnostik	77
2.3.3.3 Qualitative Merkmale politikdidaktischer Kompetenzen	82
2.3.4 Politikdidaktisches Kompetenzen im fachdidaktischen Diskurs	94
2.4 Politikdidaktische Überlegungen zur weiteren Modell- und Instrumentenentwicklung	106
3 Erweiterte Modellierung professionellen Wissens und Könnens von Politiklehrkräften	111
3.1 Ein erweitertes mehrdimensionales Modell des Professionswissens von Politiklehrkräften	111
3.2 Konzeptualisierung des politikdidaktischen Wissens und Könnens von Lehrkräften	115
3.2.1 Konzeptualisierung der Strukturdimension <i>Facetten</i>	116
3.2.2 Konzeptualisierung der Strukturdimension <i>Inhaltsbereiche</i>	117
3.2.3 Konzeptualisierung der Strukturdimension <i>Wissensarten</i>	118

4 Empirische Evidenzen	120
5 Forschungsfragen	126
6 Design der Studie	128
6.1 Anlage der Studie	128
6.2 Testablauf	129
6.3 Entwicklung der Testinstrumente	131
6.3.1 Operationalisierung der Konstrukte	134
6.3.1.1 Entwicklung der Test-Items zum Fachwissen	135
6.3.1.2 Entwicklung der Test-Items zum politikdidaktischen Wissen	136
6.3.2 Beurteilungskategorien und Ratingverfahren	144
6.3.2.1 Entwicklung der Ratingkategorien und des Ratingmanuals	145
6.3.2.2 Ratingverfahren	148
7 Ergebnisse der Studie	151
7.1 Stichprobe	151
7.2 Empirische Befunde der Studie	160
7.2.1 Testdauer und fehlende Werte	160
7.2.2 Urteilsübereinstimmung der Rater	162
7.2.3 Itemanalyse	163
7.2.3.1 Itemschwierigkeit	164
7.2.3.2 Itemtrennschärfe	166
7.2.4 Interne Konsistenz der Skalen	171
7.2.5 Testschwierigkeit und Testgesamtleistungen	174
7.2.6 Korrelationen	179
7.3 Zusammenfassung und kritische Reflexion	185
8 Diskussion und Ausblick	193
8.1 Zusammenfassung und Schlussfolgerungen	193
8.2 Ausblick	201
9 Literaturverzeichnis	204
10 Tabellen	229
11 Abbildungen	232
Anhang	236

1 Einleitung

Das professionelle Wissen und Können von Lehrkräften ist in den letzten Jahren deutlich in den Fokus der Lehrerbildung und -forschung¹ getreten. Hintergrund hierfür ist eine steigende Anzahl empirischer Befunde, die eine hohe Qualität des Unterrichts in wesentlichen Teilen der Expertise von Lehrenden zuschreiben (Helmke, 2014; Hill, Rowan & Loewenberg Ball, 2005; Kunter & Baumert, 2011; Reusser, 2009). Dabei bilden „[d]ie Ressourcen einer Lehrperson ... einen erklärenden Kausalfaktor für Lerneffekte unter vielen“ (Vogelsang & Reinhold, 2013, S. 105), weshalb davon ausgegangen wird, dass Unterricht von weiteren kontextspezifischen Faktoren – zu denen nicht zuletzt auch Hintergrundvariablen aufseiten der Schüler/-innen sowie des Unterrichts gehören – geprägt ist (Helmke, 2014, S. 69ff.). Jedoch wird gleichzeitig betont, dass Inhalte sich nicht selbst unterrichten können, sondern so transformiert und angeboten werden müssen, dass bei den Schülerinnen und Schülern Lernaktivitäten induziert werden (Baumert & Kunter, 2011a, S. 30). Insofern hängt „der Lernerfolg ... einmal von der Qualität der Lerngelegenheiten, die bereits eine Ko-Konstruktion von Lehrkräften und Schülerinnen und Schülern darstellen, und zum anderen vom mentalen Engagement der Lernenden ab“ (Baumert et al., 2011, S. 13). Da Schüler/-innen in diesem Sinne also nicht ‚belehrt‘ werden können, müssen Lehrer/-innen „Lerngelegenheiten in einem stabilen Ordnungsrahmen variationsreich, kognitiv herausfordernd und motivierend zugleich ... gestalten“ (ebd., S. 14). Die entsprechenden kognitiven Ressourcen, über die Lehrkräfte verfügen müssen, um die hiermit verbundenen Aufgaben dauerhaft professionell bewältigen zu können, werden im Anschluss an die Lehrerexpertiseforschung (Besser & Krauss, 2009; Bromme, 1992/2014; Weinert, Schrader & Helmke, 1990) in jüngeren Ansätzen häufig unter dem von Baumert und Kunter (2006) geprägten Begriff der *professionellen Handlungskompetenz* zusammengefasst (Vogelsang & Reinhold, 2013, S. 105).

Obwohl relevante Begriffe in der Literatur teilweise uneinheitlich gebraucht werden, was die Bestimmung einer gemeinhin akzeptierten Definition erschweren kann (Frey, 2014, S. 712), werden in der Lehrerbildungsforschung Kompetenzdefinitionen in Anlehnung an Weinert (2001, S. 27f.) meist mit der Bedeutung *Fähigkeiten*, *Fertigkeiten* und *Bereitschaften* ver-

¹ Die vorliegende Arbeit bemüht sich um eine geschlechtergerechte Sprache. Wo möglich werden geschlechtsneutrale Formulierungen verwendet oder es werden beide Genera ausgeschrieben. In Komposita oder Fällen, wenn das Maskulinum eine eigene Flexionsendung hat, wird aus Gründen der Lesefreundlichkeit die einfache Form verwendet, wobei auch dann stets beide Geschlechter angesprochen sind.

wendet (vgl. auch Vogelsang & Reinhold, 2013, S. 105). Klieme et al. (2007/2009) formulieren dieses kognitionspsychologisch geprägte Kompetenzverständnis so:

„In Übereinstimmung mit Weinert (2001, S. 27f.) verstehen wir unter Kompetenzen die bei Individuen verfügbaren oder von ihnen erlernbaren kognitiven Fähigkeiten und Fertigkeiten, bestimmte Probleme zu lösen, sowie die damit verbundenen motivationalen, volitionalen und sozialen Bereitschaften und Fähigkeiten, die Problemlösungen in variablen Situationen erfolgreich und verantwortungsvoll nutzen zu können.

Kompetenz ist nach diesem Verständnis eine Disposition, die Personen befähigt, bestimmte Arten von Problemen erfolgreich zu lösen, also konkrete Anforderungssituationen eines bestimmten Typs zu bewältigen.“ (S. 72)

Eine solche Unterscheidung, „wonach professionelle Handlungskompetenz aus Professionswissen einerseits und motivationalen, volitionalen und sozialen Bereitschaften und Fähigkeiten andererseits besteht und ausgehend von grundlegenden Handlungsanforderungen einer Domäne modelliert werden kann“ (Riese & Reinhold, 2010, S. 169), liegt in anderen Domänen Studien wie *COACTIV* (*Cognitive Activation in the Classroom*) (Kunter et al., 2011) und *MT21* (*Mathematics Teaching in the 21st Century*) (Blömeke, Kaiser & Lehmann, 2008), aber auch dem politikdidaktischen Forschungsprogramm *PKP* (*Professionelle Kompetenz von [angehenden] Politiklehrer/-innen*) (Oberle, Weschenfelder & Weißeno, 2012) zugrunde.

Um Lehrerkompetenzen der empirischen Forschung zugänglich machen zu können, bedarf es einer Vorstellung davon, wie Kompetenzen strukturiert sind und erworben werden. Daher werden zur „Systematisierung von Kompetenzkomponenten, -facetten und -stufen, die für die erfolgreiche Bewältigung komplexer Aufgaben oder Anforderungen bei bestimmten Referenzgruppen als bedeutsam angesehen werden können“ (Frey, 2014, S. 713), entsprechend umfangreiche holistische *Kompetenzmodelle* entwickelt. Diese beschreiben die Binnenstrukturen der einzelnen Kompetenzaspekte und deren Verhältnis zueinander oder sie zeigen Niveaustufen und deren Sequenzierung auf (Fleischer, Koeppen, Kenk, Klieme & Leutner, 2013, S. 8f.). Mit der Kompetenzmodellierung geht in der Regel der Anspruch einher, die Kompetenzen in ihren Dimensionen und Facetten diagnostizieren zu können (Klieme, Maag-Merki & Hartig, 2007, S. 8ff.), bspw. in Form von Leistungstests (Maag Merki & Werner, 2014, S. 751).

In solchen Tests erhält das *Professionswissen* von Lehrkräften als eine zentrale kognitive Ressource professioneller Handlungskompetenzen häufig ein besonderes Gewicht (Baumert et al., 2011, S. 14), weil davon ausgegangen wird, dass eine solide Wissensbasis Grundlage für erfolgreiches berufliches Handeln ist (Blömeke et al., 2014, S. 512). Die Taxonomie von Shulman (1986, 1987) bildet dabei zur Präzisierung des Wissensbegriffs einen – in der deutschsprachigen Forschung weitgehend angenommenen – theoretischen Referenzrahmen (Weißen, Weschenfelder & Oberle, 2013, S. 190). Er unterscheidet verschiedene Wissenskategorien, von denen das Fachwissen (*content knowledge [CK]*), das fachdidaktische Wissen (*pedagogical content knowledge [PCK]*) und das pädagogische Wissen (*pedagogical knowledge [PK]*) am häufigsten rezipiert werden (ebd.). Die Bedeutung des *fachdidaktischen Wissens* für das Verständnis des Wissens von Lehrkräften wird dabei hervorgehoben, da es diejenigen Facetten beschreibt, über die diese bei der Bewältigung unterrichtlicher Anforderungssituationen verfügen müssen (Gramzow, Riese & Reinhold, 2013, S. 8; Schmelzing, Wüsten, Sandmann & Neuhaus, 2010, S. 191; Shulman, 1987, S. 8; Tepner et al., 2012, S. 13). Nach Shulman (1986, 1987) wird hierunter unter anderem „das Wissen über Erklären und Darstellen“ (Krauss et al., 2008b, S. 228) sowie das Wissen „über fachbezogene Schülerkognitionen“ (ebd.) subsumiert. Die Formulierungen deuten bereits an, dass ein breites Verständnis des Wissensbegriffs angelegt wird, wonach das Vorhandensein theoretisch-formalen Professionswissens zwar eine notwendige Voraussetzung für kompetentes Handeln ist, Lehrkräfte jedoch auch über Wissen darüber verfügen müssen, wie sie dieses innerhalb beruflicher Anforderungssituationen in eine adäquate unterrichtliche Performanz überführen können. Kompetenz kann „[i]n diesem Sinne ... verstanden werden als die Verbindung von Wissen und Können in der Bewältigung von Handlungsanforderungen“ (Klieme & Hartig, 2007, S. 19).

Die Begriffspaarung *Wissen und Können* wird in der Lehrerkompetenzforschung häufig in Anlehnung an kognitionspsychologische Theorien (vgl. u. a. Anderson, 1983) verwendet, um verschiedene *Arten* oder *Typen* mental repräsentierten Wissens von Lehrkräften zu beschreiben (vgl. u. a. Lindmeier, 2011; Schmelzing et al., 2010; Tepner et al., 2012). Dabei wird in der Regel zwischen deklarativem und prozeduralen Wissen unterschieden (Gramzow et al., 2013, S. 16; Klauer & Leutner, 2012, S. 28). Das deklarative Wissen bezieht sich auf „die Kenntnis von Konzepten und Fakten“ (Riese & Reinhold, 2010, S. 172) (Wissen, dass ...), während „[u]nter dem prozeduralen Wissen ... die Kenntnis von

(Unterrichts-)Handlungen und -prozessen als dem *Wissen, wie* etwas abläuft, zu fassen [ist]“ (Tepner et al., 2012, S. 17). Insbesondere im Verständnis über das prozedurale Wissen zeichnen sich unterschiedliche wissenschaftstheoretische Positionen ab (Gramzow et al., 2013, S. 18).

Ungeachtet der teilweise differierenden theoretischen Bezugspunkte unterschiedlicher Konzeptualisierungen von Lehrerkompetenzen werden durchgängig die Herausforderungen einer adäquaten Kompetenzmessung hervorgehoben, sofern Kompetenzen als Einheit aus Wissen und Können (Baumert & Kunter, 2006, S. 481) aufgefasst werden (Frey, 2014, S. 713). Klieme et al. (2007/2009) betonen, „dass die Erfassung von Kompetenz einen breiten Begriff von Aufgaben bzw. Tests erfordert, der nicht mit Wissensabfragen allein erschöpft ist“ (S. 72). Doch obwohl sich „[n]ahezu alle Autoren und Forschungsansätze ... darin einig [sind], dass der Begriff ‚Kompetenz‘ sich nicht allein über *zugrundeliegendes Wissen und vorhandene Fertigkeiten* definiert, sondern *deren erfolgreiche Anwendung* einschließt“ (Frey, 2014, S. 713), werden in Forschungsprogrammen zur Erhebung professioneller Kompetenzen von Lehrkräften (vgl. u. a. Klieme & Leutner, 2006; Kunter et al., 2011) Diagnoseinstrumente häufig in Form von „Wissenstests [eingesetzt], die aus testökonomischen Gründen meist schriftlich realisiert werden“ (Vogelsang & Reinhold, 2013, S. 104). Das damit verbundene theoretische Verständnis zur angenommenen Handlungswirksamkeit des so operationalisierten Wissens wird ebenso problematisiert (Neuweg, 2014; Vogelsang & Reinhold, 2013) wie die entsprechenden forschungsmethodischen Zugänge, die häufig zur Anwendung kommen (Frey & Hartig, 2013; Oser, Heinzer & Salzmann, 2010, S. 5f.; Riese & Reinhold, 2010, S. 172).

Kritisch diskutiert wird hierbei, ob ein Leistungstest, bei dem Wissen als „kognitive Disposition“ (Vogelsang & Reinhold, 2013, S. 105; vgl. auch Fleischer et al., 2013, S. 6f.) im Sinne einer Ressource für professionelles unterrichtliches Handeln durch einen traditionellen schriftlichen Test erhoben werden soll, auch tatsächlich handlungsnahes Wissen erfassen kann (Vogelsang & Reinhold, 2013, S. 106). Denn es ist fraglich, ob „[d]as eigentlich Zentrale des Lehrerberufs als Profession – nämlich die Situativität, Authentizität, Komplexität und die Kontextgebundenheit des unterrichtlichen Handelns“ (Oser et al., 2010, S. 6) in schriftlichen Tests abbildbar ist, weshalb vor allem die „Validität der Maße hinterfragt [wird]“ (Lindmeier, 2013, S. 52). Stattdessen werden unterschiedliche Formate videogestütz-

ter Forschung vorgeschlagen (Blömeke, 2013b, S. 35ff.), denen das Potenzial zugeschrieben wird, „die professionellen Anforderungen des Unterrichtens stärker zu berücksichtigen“ (Lindmeier, 2013, S. 52), sodass „mit ihnen die Erhebung von professionellem Wissen und Fähigkeiten möglichst unterrichtsnah situiert werden [kann] (Kersting, 2008; Beck et al., 2008; Lindmeier, 2011)“ (ebd., S. 54). Zwar nahm die Anzahl der Studien, die mithilfe videogestützter Methoden die Erfassung von Lehrerkompetenzen fokussieren, in den letzten Jahren deutlich zu (vgl. u. a. Blömeke et al., 2014; König et al., 2014; Knievel, Lindmeier & Heinze, 2015; Lindmeier, Heinze & Reiss, 2012; Oser et al., 2010), jedoch wird nach wie vor der Forschungsbedarf zur (Weiter-)Entwicklung *standardisierter* Messinstrumente betont (Lindmeier, 2013, S. 57).

Ein solcher Forschungsbedarf lässt sich aus domänenspezifischer Perspektive auch für die Politikdidaktik formulieren (Weißeno et al., 2013, S. 200), was vor allem darin begründet sein dürfte, dass eine systematische Auseinandersetzung mit der Entwicklung von Kompetenzmodellen sowie mit Ansätzen einer standardisierten Erhebung erst seit einigen Jahren erfolgt. Die hierzu bislang in der Politikdidaktik gering ausgeprägte Forschungstradition ist auch insoweit verwunderlich, als es im Zuge der zunehmenden Orientierung an dem *Outcome* von Unterricht bereits eine intensive Auseinandersetzung mit domänenspezifischen Bildungsstandards und Schülerkompetenzmodellen gibt. So kann die Politikdidaktik zwar auf elaborierte Modellierungen der Schülerkompetenzen zurückgreifen (Detjen, Massing, Richter & Weißeno, 2012), jedoch hat erst das PKP-Forschungsprogramm (Oberle et al., 2012) eine kompetenztheoretisch fundierte domänenspezifische Auseinandersetzung mit den Kognitionen von Lehrkräften – normative Vorstellungen zum Wissen von Politiklehrenden gab es auch schon davor – angestoßen. Die PKP-Studie konnte bereits wichtige Fragen zur Struktur und zum Ausprägungsniveau des professionellen Wissens und der professionellen Überzeugungen von Politiklehrkräften, Referendarinnen und Referendaren sowie Studierenden beantworten. Gleichzeitig werden weiterhin Forschungspotenziale – unter anderem zum Einsatz von videogestützten Forschungsmethoden – für die Politikdidaktik erkannt, um die Komplexität und den Anforderungsreichtum unterrichtlicher Lehr-Lern-Situationen in Erhebungsverfahren noch stärker berücksichtigen zu können (Weschenfelder, 2014a, S. 173f).

Die vorliegende Arbeit möchte dies zum Anlass nehmen, die Frage- und Problemstellungen im Zusammenhang mit der Modellierung und Operationalisierung handlungsnaher Aspekte

professioneller Kompetenzen von Politiklehrkräften für die Politikdidaktik theoretisch und empirisch zu explorieren, um hierdurch weitere Hypothesen zu generieren und Ansätze für eine standardisierte Diagnostik aufzeigen zu können. Der Fokus liegt dabei auf dem für das Lehrerhandeln als wichtig erachteten fachdidaktischen Wissen (Krauss et al., 2008b; Schmelzing et al., 2010).

Hierzu werden im *ersten Teil* der Arbeit (*theoriebasierte Exploration*) zunächst die Begriffe des Wissens, Könnens und Handelns von Lehrkräften, wie sie in der Expertiseforschung verwendet werden, analysiert. Anschließend wird die Taxonomie des Lehrerwissens nach Shulman (1986, 1987) als Ausgangslage für die Betrachtung vorliegender Konzeptualisierungen des fachdidaktischen Wissens verwendet; Ansätze aus Nachbardomänen werden hierbei als Bezugspunkte für eine domänenspezifische Konzeptualisierung des politikdidaktischen Wissens genutzt. Wie bereits einleitend angesprochen, ist das fachdidaktische Wissen kein isolierter Aspekt der Expertise von Lehrkräften, sondern wird im Rahmen holistischer Modelle dargestellt, weshalb im Folgenden der Fokus auf der Modellierung professioneller Handlungskompetenzen liegt. Hierzu wird zunächst der kognitionspsychologisch geprägte Kompetenzbegriff erläutert. Dieser wird dann als Grundlage für die Darstellung von theoretischen Modellierungsansätzen und methodischen Diagnosemöglichkeiten genutzt. Es wird dabei erläutert, welche Ansätze einer standardisierten Erhebung handlungsnaher Kognitionen von Lehrkräften für die Politikdidaktik vorstellbar sind. Zu diesem Zweck werden vor allem Möglichkeiten und Grenzen computerbasierter videogestützter Verfahren dargestellt, denen das Potenzial zugeschrieben wird, handlungsbezogenes Wissen valider erfassen zu können als schriftliche Tests (Blömeke, 2013b, S. 35ff.; Frey & Hartig, 2013; Lindmeier, 2013). In diesem Zusammenhang werden auch Gütemaßstäbe, die in videogestützten Verfahren mit eher offenen Antwortformaten als Referenzpunkte in kompetenzdiagnostischen Verfahren heranzuziehen sind, vor dem Hintergrund domänenspezifischer Überlegungen diskutiert. Diese theoretischen Überlegungen bilden im ersten Teil der vorliegenden Arbeit schließlich die Ausgangsbasis für den Vorschlag einer erweiterten mehrdimensionalen Modellierung des professionellen Wissens und Könnens, die sich einerseits anschlussfähig an bisherige generische und domänenspezifische Modelle zeigt, andererseits jedoch auch den Herausforderungen einer umfassenderen Konzeptualisierung handlungsnaher Aspekte professioneller Lehrerkompetenzen gerecht werden soll.

Im *zweiten Teil* der der Untersuchung (*empirische Exploration*) wird diese erweiterte Modellierung des politikdidaktischen Wissens und Könnens von Politiklehrkräften in ein erkundendes Testverfahren umgesetzt, das vorläufige Aussagen zu den Strukturannahmen des Modells sowie zu den Möglichkeiten einer computerbasierten videogestützten Diagnostik ermöglichen soll. Dieser empirisch-explorative Teil der Arbeit dient somit der Instrumenten(weiter)entwicklung sowie der Hypothesengenerierung. Hierzu werden angehenden Politiklehrkräften im Format des *video-cued testing* Videovignetten mit authentischen, herausfordernden und anforderungsreichen Unterrichtssituationen gezeigt, auf die sie in natürlicher Sprache unter Zeitdruck reagieren müssen. Im Unterschied zu videobasierten Verfahren, die beispielsweise versuchen, aus der Beobachtung metakommunikativ-reflektierender Reaktionen der Probanden Rückschlüsse auf die dahinterliegenden Kompetenzen zu ziehen, ist das Design der vorliegenden Studie als Leistungstest ausgelegt. Die Testpersonen verhalten sich in diesem Verfahren verbal so, als wenn sie im Unterricht wären; es wird davon ausgegangen, dass ihre Fähigkeit nahe an authentischen Unterrichtssituationen geschätzt werden können (Lindmeier, 2011; Lindmeier et al., 2012).

Die empirischen Befunde der vorliegenden explorativen Studie deuten dabei darauf hin, dass es möglich ist, das professionelle politikdidaktische Wissen und Können in dem für diese Arbeit zu konzeptualisierenden Modell abzubilden. So korrelieren die Testleistungen der Probanden in den für die Erhebung gebildeten Skalen zum Fachwissen, zum deklarativen politikdidaktischen Wissen und zum prozeduralen politikdidaktischen Wissen moderat, was als erstes – wegen des explorativen Charakters der Studie zunächst probeweise zu formulierendes – Anzeichen dafür interpretiert werden kann, dass die Skalen kein generelles Konstrukt erfassen. Insbesondere die Skala zum prozeduralen Wissen, die mithilfe von Videovignetten erfasst wird, unterscheidet sich deutlich von den beiden anderen Skalen. Der Vergleich der Testleistungen nach Gruppen zeigt im Wesentlichen kongruente Ergebnisse unter anderem zur PKP-Studie; so liegen beispielsweise die Leistungen der Sekundarstufe-II-Studierenden auf höherem Niveau als die Testleistungen der Sekundarstufe-I-Studierenden. Die vorläufigen empirischen Befunde werden abschließend vor dem Hintergrund der zugrunde liegenden theoretischen Annahmen der Arbeit kritisch reflektiert und zusätzlich werden Implikationen für zukünftige Forschungen und die Lehrerbildung diskutiert.

2 Politikdidaktisches Wissen und Können von Lehrkräften

In diesem Kapitel werden zunächst die für die vorliegende Arbeit zentralen Begriffe des Wissens, Könnens und Handelns von Lehrkräften analysiert (Kapitel 2.1). Hierzu wird von einem Expertisebegriff ausgegangen, nach dem das professionelle Wissen eine kognitive Ressource für professionelles Unterrichtshandeln ist (Blömeke et al., 2014, S. 512). Dieses Wissen wird im weiteren Verlauf des Kapitels in Anlehnung an die theoretischen Arbeiten von Shulman (1986, 1987) in verschiedene Kategorien taxonomisch untergliedert (Kapitel 2.1.1) und es werden ausgewählte Vorstellungen zu seiner mentalen Repräsentation erläutert (Kapitel 2.1.2). In Kapitel 2.2 steht die Frage im Mittelpunkt, wie sich das für die professionelle Berufsausübung zentrale politikdidaktische Wissen von Lehrkräften theoretisch darstellen lässt. Hierzu werden unterschiedliche Konzeptualisierungen des fachdidaktischen Wissens in ihrer internen Strukturierung sowie den ihnen zugrunde liegenden Ansätzen zur Operationalisierung der konzeptualisierten Konstrukte miteinander verglichen, um so identifizieren zu können, welche Gemeinsamkeiten und Unterschiede sich in den Ansätzen abzeichnen. In Kapitel 2.3 wird anschließend vor dem Hintergrund eines kognitionspsychologischen Kompetenzbegriffs (Kapitel 2.3.1) die Verortung des fachdidaktischen Wissens in übergeordneten holistischen Modellen professioneller Kompetenzen erläutert (Kapitel 2.3.2). Die Herausforderungen einer domänenspezifischen Kompetenzdiagnostik werden in Kapitel 2.3.3 reflektiert. In Kapitel 2.3.4 erfolgt eine Gegenüberstellung von unterschiedlichen Modellierungsansätzen professioneller Handlungskompetenzen von Politiklehrkräften, bevor in Kapitel 2.4 abschließende Überlegungen zur weiteren Modell- und Instrumentenentwicklung aus einer politikdidaktischen Perspektive vorgestellt werden.

2.1 Zum professionellen Wissen, Können und Handeln von Lehrkräften

Beschreibt man Unterricht als die zentrale Aufgabe von Lehrkräften (Blömeke, 2002, S. 42; Baumert et al., 2011, S. 9f.), so erscheint es zunächst intuitiv nachvollziehbar, dass sie *Wissen* darüber haben müssen, wie Unterricht ermöglicht werden kann (Tenorth, 2006, S. 585). Die Lehrerbildungsforschung beschäftigt sich jedoch erst seit den 1980er-Jahren explizit mit diesem Wissen in seinem heutigen Verständnis (Neuweg, 2014, S. 583). Das in der ersten Hälfte des 20. Jahrhunderts breit rezipierte *Persönlichkeitsparadigma* ging zunächst davon aus, dass positive Merkmale von Lehrerpersönlichkeit (Engagement, Freundlichkeit, Zuge-

wandtheit) ein oder sogar *der* bestimmende Prädiktor für die Lernleistungen der Schülerinnen und Schüler ist, obgleich Versuche, ebensolche Eigenschaften von Lehrkräften auch empirisch zu bestimmen, als nicht zielführend gelten (Helmke, 2014, S. 46).² Demgegenüber stellt das in den 1960er-Jahren aufkommende *Prozess-Produkt-Paradigma* Merkmale in den Vordergrund, die mit bestimmten Zielvariablen des Unterrichts in Beziehung gesetzt werden. Hierbei besteht „[d]ie Methodik des Prozess-Produkt-Paradigmas ... aus der empirischen Erfassung bestimmter Aspekte des Unterrichts (sogenannte ‚Prozesse‘, z. B. Bestimmung der Anzahl anspruchsvoller Lehrerfragen pro Zeiteinheit), der Erfassung von Zielkriterien des Unterrichts (sogenannte ‚Produkte‘, z. B. des Lernzuwachses) und der Berechnung von Zusammenhangsmaßen (etwa Korrelationen) zwischen den untersuchten Prozessen und Produkten“ (Krauss & Bruckmaier, 2014, S. 241). Die eingangs angesprochene Fokussierung des Lehrerwissens vollzog sich schließlich ab den 1980er-Jahren „beeinflusst von der ‚kognitiven Wende‘ und der damit verbundenen Entstehung der allgemeinen Expertiseforschung in der kognitiven Psychologie“ (Besser, 2014, S. 14). Beim *Expertenparadigma* wird, „[w]ie schon beim Persönlichkeitsparadigma[,] ... wieder nach dem ‚guten Lehrer‘ gesucht, diesmal aber stehen nicht mehr Charaktereigenschaften oder Persönlichkeitsmerkmale im Vordergrund, sondern vielmehr Wissen und Können des Lehrers im Sinne eines ‚kompetenten Fachmanns für das Unterrichten‘ (Helmke, 2003)“ (Krauss & Bruckmaier, 2014, S. 242). Es wird hierbei angenommen, dass die Expertise von Lehrkräften nicht an ein angeborenes Talent gebunden ist, sondern auch durch zielgerichtete Übung aufgebaut werden kann (Ericsson, Krampe & Tesch-Romer, 1993; vgl. auch Anderson, 2013, S. 209).³

Bei der Darstellung von Expertiseansätzen ist zu berücksichtigen, dass grundlegende Begriffe unterschiedlich konnotiert sind (Besser, 2014, S. 15). So kann *Expertise* „einmal das Wissen eines Experten, einmal die Fähigkeit zu überdurchschnittlichen Leistungen und einmal die permanente Erbringung von Höchstleistungen [bezeichnen]“ (Krauss & Bruckmaier, 2014, S. 244). Ebenso ist eine eindeutige Definition des Begriffs *Experte* in Bezug auf den Lehrerberuf schwierig, was Besser (2014) zusammenfassend so problematisiert:

² Für eine ausführliche Darstellung des Persönlichkeitsansatzes in der Unterrichtsforschung, der in der vorliegenden Arbeit nicht umfassend referiert werden kann, vgl. Mayr (2014).

³ Bei dieser Form aktiven Lernens „erhalten [die Lernenden] Feedback über ihre Leistung und registrieren sorgfältig, wie gut ihre eigene Leistung der korrekten Leistung entspricht und wo sie davon abweichen“ (Anderson, 2013, S. 209).

- „Wann wird man wohl die Leistung eines Lehrers als ‚herausragend‘ bezeichnen?
- Was genau kann als ‚das Gebiet‘ angesehen werden, auf dem ein Lehrer eine geforderte, herausragende Leistung erbringen soll?
- Wie kann letztendlich die Leistungsstärke eines Lehrers tatsächlich gemessen bzw. ermittelt werden, wenn nicht einmal klar zu sein scheint, auf welchem speziellen Gebiet eine Lehrkraft welche Leistung zu erbringen hat?“ (S. 16f.)

Daher wird für die Möglichkeiten der Verwendung eines *leistungsbasierten Expertisebegriffs* in der Lehrerforschung darauf hingewiesen, dass „[d]ie Abkehr vom Höchstleistungsgedanken zur Identifikation von Experten ... aus wissenschaftlicher Sicht vor allem in *schlechtdefinierten* Domänen statt[findet]“ (Krauss & Bruckmeier, 2014, S. 247). Hinzu kommt, dass es für Forschungsansätze im Sinne eines *Experten-Novizen-Paradigmas*, bei dem kontrastiv Experten mit Anfängern verglichen werden (Blömeke, 2002, S. 77; Weinert et al., 1990; S. 176ff.), „kein reliables und valides Expertisemaß [gibt]“ (Krauss & Bruckmeier, 2014, S. 249), insofern also unklar ist, was Expertise im Lehramt auszeichnet. Da es wegen konkurrierender Schülerkompetenzmodelle, unterschiedlicher Schulfachkonzeptionen (zum Teil schon innerhalb eines Bundeslandes) und – trotz des Übergangs zu Kernlehrplänen – fehlender domänenspezifischer Bildungsstandards keine sich von den etablierten allgemeinen Qualitätsmerkmalen des Schulunterrichts (Helmke, 2014) abhebenden domänenspezifischen Zielleistungskriterien für Politikunterricht gibt (vgl. Kapitel 2.2.1), die sich einer empirischen Messbarkeit zugänglich zeigen, dürfte eine Orientierung an einem leistungsbasierten Expertisebegriff für diese Domäne nur schwer möglich sein.

Stattdessen scheint die Zugrundelegung eines *wissensbasierten Expertisebegriffs*, bei dem „[d]urch eine möglichst umfassende Rekonstruktion der Arbeitssituation ... Rückschlüsse auf das Wissen gezogen werden [soll], das zur Bewältigung der Anforderungen ... (z.B. berufliche Zielvorstellungen, Sachgegebenheiten und Bedingungen der Problembearbeitung) notwendig ist“ (Krauss & Bruckmaier, 2014, S. 247f.), für eine solche gering strukturierte Domäne zielführender. Bromme (1992/2014), der durch seine theoretische Arbeiten „die Bezeichnung ‚Experten‘ für die Mitglieder der Lehrerprofession entscheidend mitprägte“ (Krauss & Bruckmaier, 2014, S. 251), betont, dass „[d]ie Untersuchung des Expertendenkens und -wissens ... die parallele Analyse der Anforderungen an den Experten [erfordert]“ (Bromme, 1992/2014, S. 39) oder wie es Lindmeier (2011) beispielhaft für die Mathematik formuliert: „It may be reasonable to conjecture that mathematics teachers‘ knowledge and

competencies are organized around typical teacher tasks, i.e. around tasks that frequently recur as professional demands“ (S. 88).

Die Analyse von typischen Aufgaben, die Lehrer/-innen im Berufsalltag zu bewältigen haben, führt Bromme (1992/2014) zur Benennung von drei *Grundanforderungen des Unterrichts*: „a) Die Organisation und Aufrechterhaltung einer Struktur von Lehrer- und Schüleraktivitäten, b) die Entwicklung des fachbezogenen Wissens und c) die Aufteilung von Unterrichtszeit und die Steuerung des Unterrichtstempos“ (S. 76). Zu berücksichtigen ist, dass der Lehrerberuf grundsätzlich durch Unsicherheiten geprägt ist, was eine Rekonstruktion beruflicher Anforderungssituationen erschwert (Weschenfelder, 2014a, S. 31f.). Baumert und Kunter (2006) beschreiben diese „Erfolgsunsicherheit des Lehrerhandelns“ (S. 476) in einem doppelten Sinne:

„Dies gilt einmal – wie TENORTH ... herausstellt – in einem generellen Sinne, insofern die Berufsaufgabe von Professionellen weder normativ noch operativ eindeutig vorgegeben ist, sondern unter Berücksichtigung der jeweiligen Umstände des Einzelfalls und der jeweiligen Situation ausdeutungs- und konkretionsbedürftig ist. Dies gilt aber auch in einem spezifischen Sinn, insofern es keine technischen Transformationsregeln gibt, die dem Handelnden Erfolg garantieren. Erfolg stellt sich nicht unabhängig vom Handeln der Klienten ein.“ (Ebd.)

Der hiermit angesprochene Anforderungsreichtum des Schulunterrichts kann in einem *Angebots-Nutzungs-Modell* (Helmke, 2014, S. 69ff.) genauer beschrieben werden (Abbildung 1).

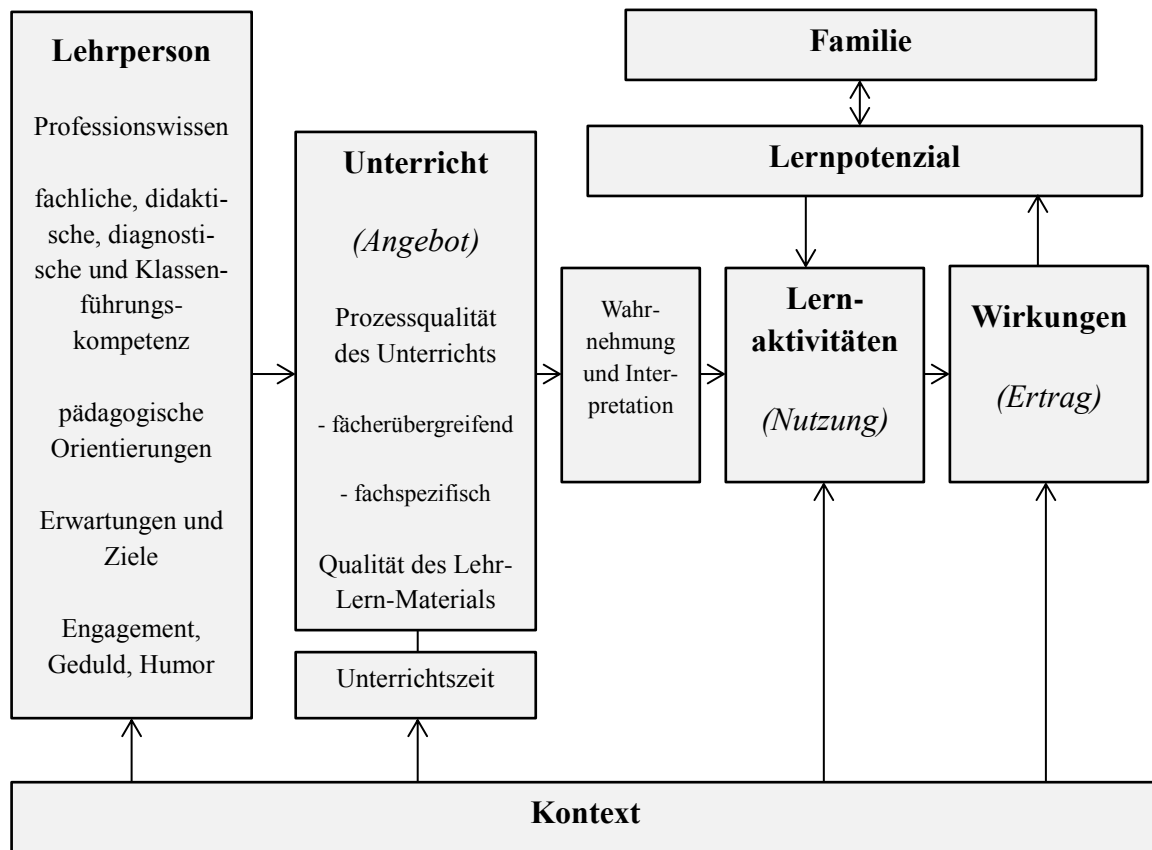


Abbildung 1: Angebots-Nutzungs-Modell der Wirkungsweise des Unterrichts (eigene Darstellung in Anlehnung an Helmke, 2014, S. 71)

Die Nutzung des Lernangebotes hängt diesem Modell zufolge nicht allein von der durch die Lehrkraft beeinflussbaren Qualität dieses Angebotes ab, sondern auch von einer Reihe Hintergrundvariablen, die die Lehrperson nicht – oder zumindest nicht *direkt* – planen und steuern kann. Dazu gehört unter anderem der Kontext des Unterrichts (Schulform, Klassenzusammensetzung), das Lernpotenzial der Schülerinnen und Schüler (allgemeine kognitive Fähigkeiten, Vorkenntnisse sowie Motivation und Selbstwirksamkeitsüberzeugungen) sowie der familiäre Hintergrund der Lernenden (ebd.). Insofern führt „[d]as unterrichtliche Angebot ... nicht notwendigerweise direkt zu den *Wirkungen* (äußerer rechter Kasten ...), sondern seine Wirksamkeit für das Lernen hängt von zweierlei Typen von vermittelnden Prozessen auf Schülerseite ab: (1) davon, ob und wie Erwartungen überhaupt *wahrgenommen* und wie sie *interpretiert* werden, sowie (2) ob und zu welchen motivationalen, emotionalen und volitionalen (auf den Willen bezogenen) Prozessen sie auf der Schülerseite führen“ (ebd., S. 71).

Den hiermit skizzierten ko-konstruktiven Charakter von Unterricht (Baumert & Kunter, 2006, S. 477) beschreibt Bromme (1992/2014) so:

„Der Lehrer kann also den Unterrichtsprozeß nur in einer stochastischen, nicht in einer deterministischen Weise beeinflussen. Außerdem sind seine Handlungen ein Teil des Systems ‘Unterricht’. Dadurch hat er einen besonderen Status: Er ist einerseits selber Teil des Systems ‘Unterricht’, andererseits muß er es zu überwachen und zu steuern versuchen.“ (S. 82)

Die Eigenschaften der Lehrperson, die für die Bewältigung dieser Aufgaben notwendig sind, werden im Modell von Helmke (2014, S. 76) zwar nicht direkt den unterrichtlichen Qualitätsmerkmalen zugeordnet, dennoch sind sie für ein gelingendes Lernen bedeutsam, da sie zum einen als Persönlichkeitsmerkmale (Geduld, Humor Engagement), aber auch in ihrer Ausprägung als fachwissenschaftliche und fachdidaktische Expertise maßgeblich Einfluss auf die Qualität des Lernangebotes haben. Helmke verwendet in diesem Zusammenhang auch den Begriff der „unterrichtsrelevante[n] Expertise“ (ebd.).

Der Ansatz einer Rekonstruktion beruflicher Anforderungssituationen, bei dem „[e]ine Analyse der Handlungsanforderungen und Handlungskompetenzen ... am Kern der Berufstätigkeit bei der Vorbereitung, Inszenierung, Durchführung und Auswertung von Unterricht anzusetzen [hat]“ (Baumert et al., 2011, S. 9), ist Ausgangslage für unterschiedliche Forschungen zum Lehrerwissen. So werden in COACTIV (Kunter et al., 2011) durch eine Analyse beruflicher Anforderungen diejenigen Lehrermerkmale zu bestimmen versucht, „die unmittelbar notwendige Voraussetzungen für die Realisierung eines qualitätsvollen Unterrichts darstellen“ (Baumert et al., 2011, S. 14). Unter Berücksichtigung der beruflichen Unsicherheiten wird ein wissensbasierter Expertisebegriff verwendet, bei dem „[a]ls Gütemaßstab ... die – wie Oser (2009) formulierte – souveräne, das heißt zuverlässige und dauerhafte Bewältigung der beruflichen Anforderungen eingeführt [wird]“ (ebd.), womit der Abgrenzung zu einem an Höchstleistungen orientierten Expertisebegriff Rechnung getragen wird (ebd.). Die Konzeptualisierungen des Lehrerwissens in der MT21-Studie (Blömeke et al., 2008) sowie in den Folgestudien *TEDS-M* (*Teacher Education and Development Study in Mathematics*) (Blömeke, Kaiser & Lehmann, 2010a; Blömeke, Kaiser & Lehmann, 2010b) und *TEDS-LT* (*Teacher Education and Development Study: Learning to Teach*) (Blömeke et al., 2011; Blömeke et al., 2013) erfolgten ebenfalls entlang eines wissensbasierten Expertisebegriffs. Es wird hierbei davon ausgegangen, dass der Beruf des Lehrers/der Lehrerin eine „ausbildungsintensive Profession dar[stellt] und ... durch komplexe berufliche

Anforderungen gekennzeichnet [ist]“ (Blömeke, 2013a, S. 10). Zwar wird das Lehrerwissen in TEDS-M nicht anhand konkreter beruflicher Anforderungssituationen rekonstruiert (Lindmeier, 2011, S. 94), jedoch werden in der TEDS-LT-Studie kognitive Anforderungen ausdifferenziert, die der Strukturierung des professionellen Wissens von Lehrkräften dienen. Demnach umfassen „[k]ognitive Anforderungen ... neben den Wissenstypen Fakten- und begriffliches Wissen, prozedurales Wissen und metakognitives Wissen auf einer zweiten Dimension die Prozesse Erinnern (*remember*), Verstehen (*understand*), Anwenden (*apply*), Analysieren (*analyze*), Bewerten (*evaluate*) und Erschaffen (*create*)“ (Blömeke, 2013a, S. 11). Solche kognitiven Prozesse sind beispielsweise das „Erinnern und Abrufen“ (König, Blömeke & Doll, 2011, S. 139) oder das „Verstehen und Anwenden“ (ebd.). Diese Aufteilung dient – in einer für die TEDS-Studien synthetisierten Form – der differenzierten Testaufgabenentwicklung. So muss beispielsweise „[z]ur Bearbeitung von Testaufgaben, die kognitive Prozesse des Verstehens und Anwendens benötigen, ... vorhandenes Wissen mit einer Problemstellung in Verbindung gebracht werden“ (ebd.).

Schließlich lässt sich, wenn die Perspektive von der Forschung auf die Bildungsadministration und -politik verschoben wird, der Ansatz einer Rekonstruktion beruflicher Anforderungssituationen zur Bestimmung der notwendigen Expertise von Lehrkräften auch in berufsbezogenen Standards erkennen.⁴ So wird deren Hauptaufgabe in den *Standards für die Lehrerbildung: Bericht der Arbeitsgruppe der Kultusministerkonferenz* (KMK, 2004) folgendermaßen beschrieben:

„1. *Lehrerinnen und Lehrer sind Fachleute für das Lehren und Lernen.* Ihre Kernaufgabe ist die gezielte und nach wissenschaftlichen Erkenntnissen gestaltete Planung, Organisation und Reflexion von Lehr- und Lernprozessen sowie ihre individuelle Bewertung und systemische Evaluation. Die berufliche Qualität von Lehrkräften entscheidet sich an der Qualität ihres Unterrichts. Sie vermitteln grundlegende Kenntnisse, Fähigkeiten und Fertigkeiten und Methoden, die es dem Einzelnen ermöglichen, selbstständig den Prozess lebenslangen Lernens zu meistern. Die Qualität einer guten Schule und die Wirksamkeit eines guten Unterrichts werden entscheidend durch die professionellen und personalen Fähigkeiten von Lehrerinnen und Lehrern geprägt. Für die berufliche Arbeit sind umfassende fachwissenschaftliche wie

⁴ Der Begriff des *Perspektivenwechsels* wird an dieser Stelle deshalb verwendet, da die in der Lehrerbildung verwendeten Standards zwar auf theoretischen Kompetenzmodellierungen basieren können, „jedoch in der Regel über die Sammlung, Bewertung und Strukturierung durch Experten und Stakeholder entwickelt [werden]“ (Frey, 2014, S. 714), so auch die hier angeführten KMK-Standards (ebd.). Außerdem fokussieren Standards häufig nicht das Wissen und die kognitiven Fertigkeiten, sondern die Performanz von Lehrkräften als ‚Produkt‘ der Kompetenz (ebd.; vgl. auch Kapitel 2.3.1.1).

auch pädagogisch-(fach)didaktische und psychologisch-diagnostische Kompetenzen sowie kommunikative und soziale Fähigkeiten erforderlich.“ (S. 5)⁵

Als Standards werden hierbei die entlang dieser Aufgaben⁶ zu formulierenden Anforderungen an die Lehrer/-innen verstanden, die gleichzeitig als verbindlicher Bildungsauftrag an die entsprechenden Institutionen (Hochschulen und Universitäten, Zentren für schulpraktische Lehrerbildung) formuliert werden (KMK, 2015, S. 3ff.). So formuliert die KMK (2004) als Anforderungen hinsichtlich der auszubildenden fachdidaktischen Kompetenzen unter anderem, dass Lehrkräfte fähig sein sollen, „fachbezogene Formen der Lerndiagnose und Lernförderung zu kennen, zu beurteilen und einzusetzen“ (S. 14) sowie „ein breites Spektrum unterschiedlicher methodischer Lehr-Lern-Formen sachgerecht einzusetzen“ (ebd.).

Bis hierher kann zusammengefasst werden, dass es eine zentrale Aufgabe für Lehrkräfte ist, „den Unterricht in eine geeignete *soziale, zeitliche* und *inhaltliche* Struktur [zu] bringen“ (Krauss & Bruckmaier, 2014, S. 251). Nach dem wissensbasierten Expertiseansatz werden durch eine Anforderungsanalyse diejenigen Aufgaben identifiziert, die Lehrer/-innen im Berufsalltag bewältigen müssen. Dies ist auch Ausgangslage von Forschungsprogrammen zum Lehrerverwissen wie beispielsweise COACTIV (Kunter et al., 2011). Anforderungs- und Aufgabenprofile lassen sich ebenfalls in administrativen Strukturvorgaben zur Lehrerbildung (KMK 2004, 2014, 2015) identifizieren. Die Unsicherheiten dieses Berufs werden in der Forschung dabei anerkannt (Baumert & Kunter, 2006) und es wird betont, dass „[d]ie Analyse der Problemstruktur ... um so schwieriger [wird], je schlechter die Probleme definiert sind“ (Bromme, 1992/2002, S. 40). Für die Politikdidaktik dürfte es daher eine gewisse Herausforderung darstellen, zu diesem Zweck eine Präzisierung der genauen beruflichen Anforderungen bereitzustellen, die sich, sobald sie die Sphäre rein normativer Zielsetzungen verlassen, mit den genannten Limitationen fehlender domänenspezifischer Standards beschäftigen muss. Wiederum alle Fächer betrifft die Herausforderung, jeweils dasjenige bereichsspe-

⁵ Indirekt wird in dieser Formulierung auch eine Wirkkette zum Zusammenhang zwischen Kompetenzen von Lehrkräften und Lernenden angenommen, wonach „Lehrkräfte innerhalb ihrer Ausbildung und während ihrer Berufstätigkeit die notwendigen Voraussetzungen erwerben [bzw. erwerben sollten], die es ihnen ermöglichen, in schulischen Lehr-Lernsituationen so zu handeln, dass ein möglichst hoher Lernerfolg auf Seiten der Schülerinnen und Schüler resultiert“ (Vogelsang & Reinhold, 2013, S. 104).

⁶ Es werden weitere Aufgabenbereiche genannt, die in diesem Zusammenhang jedoch nicht relevant sind und daher im Folgenden unberücksichtigt bleiben.

zifische Wissen zu identifizieren, das Lehrkräfte für die zentrale Aufgabe des schulischen Unterrichtens benötigen (Blömeke et al., 2014, S. 512; Lindmeier, 2013, S. 47). Dieses Wissen wird „im Falle von Professionen ... auch ‚Professionswissen‘ genannt“ (Krauss & Bruckmaier, 2014, S. 247) und soll im folgenden Kapitel genauer dargestellt werden.⁷

2.1.1 Eine Topologie des professionellen Wissens von Lehrkräften

Der Begriff des Professionswissens wird maßgeblich von den theoretischen Beiträgen von Shulman (1986, 1987) bestimmt (Baumert & Kunter, 2006, S. 479f.; Krauss et al., 2008b, S. 226; Lindmeier, 2013, S. 47). Ausgangslage für seine Überlegungen zur Struktur und Entwicklung des professionellen Wissens von Lehrkräften ist folgende Annahme: „The rhetoric regarding the knowledge base, however, rarely specifies the character of such knowledge“ (Shulman, 1987, S. 4). Shulman betont daher, dass präzise Vorstellungen zum Lehrerwissen notwendig sind, wenn dieses Wissen nach qualitativen Standards in der Lehrerbildung modelliert und schließlich auch diagnostiziert werden soll:

„If teachers are to be certified on the basis of well-grounded judgements and standards, then those standards on which a national board relies must be legitimized by three factors: they must be closely tied to the findings of scholarship in the academic disciplines that form the curriculum (such as English, physics and history) as well as those that serve as foundations for the process of education (such as psychology, sociology, or philosophy); they must possess intuitive credibility (or ‘face validity’) in the opinions of the professional community in whose interests they have been designed; and they must relate to the appropriate normative conceptions of teaching and teacher education.“ (Ebd., S. 5)

Die von Shulman (1986, 1987) eingeführten Begriffe *content knowledge*, *pedagogical knowledge* und *pedagogical content knowledge* wurden in der deutschsprachigen Forschung als *Fachwissen*, *pädagogisches Wissen* oder *pädagogisch-psychologisches Wissen* und *fachdidaktisches Wissen* übersetzt (Krauss et al., 2008b, S. 226; Tepner et al., 2012, S. 8). Da solche Kategorien „darauf ab[zielen], auf theoretischem Wege die Dimensionen des Professionswissens möglichst vollständig zu erfassen“ (Krauss, et al., 2008b, S. 226), um-

⁷ Wenn im Folgenden von *professionellem* Wissen und Können gesprochen wird, folgt dies der Auffassung, dass „[p]olitische Bildung ... als Profession bezeichnet [wird]“ (Hufer & Richter, 2013, S. 13). Hinsichtlich der Begriffe *Profession* (Berufsstatus), *Professionalisierung* (Entwicklung professionellen Wissens und Könnens) und *Professionalität* (professionelle Ausübung von Tätigkeiten) nach Huppert und Abs (2013, S. 65) steht somit letztgenannter im Fokus nachfolgender Überlegungen; für eine weitergehende begriffliche Differenzierung vgl. auch Eraut (1994/2000, S. 1ff.).

spannt das professionelle Lehrerwissen nach Shulman (1987) auch weitere Kategorien wie das „curriculum knowledge“ (S. 8), das „knowledge of educational ends, purposes, and values, and their philosophical and historical backgrounds“ (ebd.) oder das „knowledge of educational contexts, ranging from the workings of the group or classroom, the governance and financing of school districts, to the character of communities and cultures“ (ebd.), wobei eine solche umfassende Taxonomie in deutschsprachigen Ansätzen in der Regel nicht rezipiert wird (Tepner et al., 2012, S. 8f.). Das Fachwissen, das fachdidaktische Wissen und das pädagogische Wissen können daher als „die allgemein akzeptierten Kernkategorien des Professionswissens von Lehrkräften“ (Krauss & Bruckmaier, 2014, S. 251) bezeichnet werden. Unter der Annahme, dass das „Kerngeschäft [der Lehrenden] im ‚Unterricht‘ zu sehen [ist], kompetenzorientiert gedacht, also in der Fähigkeit, Lernsequenzen ergebnisbezogen zu organisieren“ (Tenorth, 2006, S. 585), wird dabei die besondere Bedeutung des Fachwissens sowie des fachdidaktischen Wissens betont (Baumert & Kunter, 2011a, S. 35) und damit gleichzeitig die Notwendigkeit einer domänenspezifischen Ausformulierung des Professionswissens hervorgehoben (Baumert & Kunter, 2006, S. 483; Krauss et al., 2008b, S. 227; vgl. auch Kapitel 2.2).

Wenn nach einem solchen Verständnis das Professionswissen als ein Teilaspekt von „kognitiven Kompetenzen“ (Krauss & Bruckmaier, 2014, S. 251) oder – allgemeiner formuliert – als „kognitive Disposition“ (Vogelsang & Reinhold, 2013, S. 105) bezeichnet wird, liegt dem meist die Annahme zugrunde, dass dieses Wissen eine Basis für die „erfolgreiche Ausübung des Lehrerberufs“ (Blömeke et al., 2014, S. 512) und daher auch eine zentrale Ressource für das unterrichtliche Handeln darstellt (Vogelsang & Reinhold, 2013, S. 105). Entsprechende Wirkvorstellungen zum Wissen finden sich in Ansätzen, die sich auf das Konzept der Expertise beziehen, ebenso wie in Forschungen, die auf Konzepte zum *Habitus* und zu *subjektiven Theorien* rekurrieren.⁸ Da innerhalb der Bandbreite entsprechender Forschungsansätze – aber auch in den einzelnen Ansätzen – die Begriffe Wissen, Können und Handeln nicht immer einheitlich verwendet werden, entsteht eine Mehrdeutigkeit des Wissensbegriffs (Neuweg, 2014, S. 583), die klärungsbedürftig ist. Hiermit angesprochen ist die Frage, wie das professionelle Wissen mental repräsentiert wird, um im Unterricht handlungswirksam werden zu

⁸ Diese können hier nicht ausführlich referiert werden; für eine Darstellung dazu, wie diese Ansätze in der Lehrerbildungsforschung rezipiert werden, vgl. Vogelsang & Reinhold, 2013.

können; hierzu wird im Folgenden der Begriff der *Wissensarten* (Gramzow et al., 2013, S. 16; Tepner et al., 2012, S. 16) verwendet.⁹

2.1.2 Zur mentalen Repräsentation von Wissen

In Expertiseansätzen wird das *Expertenwissen* (Bromme, 1992/2004, S. 38) zumeist „auch als *prozedurales Wissen* modelliert und ... als diejenige Ressource definiert, die es ermöglicht [*sic*] komplexe, kognitive Tätigkeiten durchzuführen“ (Vogelsang & Reinhold, 2013, S. 110); gleichzeitig findet hierdurch in der Regel eine Abgrenzung dieser Wissensart zum deklarativen Wissen statt.¹⁰ Aus kognitionspsychologischer Sicht werden die Unterschiede zwischen diesen Wissensarten folgendermaßen charakterisiert (Oswald & Gadenne, 1984):

„Unter deklarativem Wissen ist das Faktenwissen zu verstehen, das Personen im Gedächtnis gespeichert haben, das sie sich bewußt machen können und das sie in der Regel zu verbalisieren vermögen. Prozedurales Wissen bezieht sich auf die kognitiven Mechanismen, die Personen dazu in die Lage versetzen, komplexe, kognitive und motorische Handlungen durchzuführen, ohne dabei die einzelnen Bestandteile dieser Handlungen bewußt kontrollieren zu müssen.“ (S. 173)

Kognitionspsychologische Theorien gehen dabei davon aus, dass die beiden Wissensarten im Gedächtnis in unterschiedlicher Form repräsentiert und gespeichert werden. Das *deklarative Gedächtnis* dient in solchen Modellvorstellungen der Speicherung von „Fakten und Ereignissen“ (Zimbardo & Gerrig, 2004, S. 296) und es wird „[i]nnerhalb des deklarativen Gedächtnisses ... zwischen dem semantischen Gedächtnis für Faktenwissen und dem episodischen Gedächtnis für autobiografische Inhalte unterschieden“ (Pfeiffer, 2008, S. 77). Eine mentale Repräsentation erfolgt im deklarativen Gedächtnis jedoch nicht in einzelnen Begriffen, sondern in Konzepten, weshalb aussagegleich für das deklarative Wissen auch der Begriff des *konzeptuellen Wissens* verwendet wird (Anderson, 2013, S. 104). Da „[w]ir ... gar nicht anders [können], als die Welt in Begriffen von bekannten Kategorien zu erleben“ (ebd.), werden im Gedächtnis vorhandene Kategorien, Muster und Begriffe zu ebensolchen Konzepten zusammengefasst, was Carey (1999) wie folgt beschreibt:

⁹ Andere Untersuchungen verwenden Begriffe wie „*Wissenskomponenten*“ (Schmelzing, 2010, S. 30), „*Wissenstypen*“ (Baumert & Kunter, 2011a, S. 33; Stender, 2014, S. 20) oder im Englischen „*types of knowledge*“ (Lindmeier, 2011, S. 19)

¹⁰ Es gibt weitere Ansätze zur Systematisierung von Wissensarten; für eine Übersicht vgl. de Jong und Ferguson-Hessler (1996).

„Concepts are units of mental representation roughly the grain of single lexical items, such as *object*, *matter*, and *weight*. Beliefs are mentally represented propositions taken by the believer to be true, such as *Air is not made of matter*. Concepts are the constituents of beliefs; that is, propositions are represented by structures of concepts. Theories are complex mental structures consisting of a mentally represented domain of phenomena and explanatory principle that account for them.“ (S. 459f.)

Konzepte dienen der Vorhersagbarkeit von Situationen; gleichzeitig birgt diese Kategorisierung die Gefahr der Ausbildung von Stereotypen (Anderson, 2013, S. 104). Im Zuge des Erwerbs von Wissen ändern sich Konzepte, wobei es wissenschaftlich unterschiedliche Auffassungen dazu gibt, wie Konzepte erworben bzw. aufgebaut und ausdifferenziert werden (Carey, 1999).

Im Gegensatz zum deklarativen Gedächtnis speichert das *prozedurale Gedächtnis* Wissen darüber, „wie Dinge getan ... [und] wie perzeptuelle, kognitive und motorische Fertigkeiten erworben, aufrechterhalten und angewendet werden“ (Zimbardo & Gerrig, 2004, S. 296; vgl. auch Anderson, 2013, S. 159). Personen können in der Regel nicht ex situ formulieren, wie der ursprüngliche Lernprozess aussah (Zimbardo & Gerrig, 2004, S. 297). Das Zubinden eines Schuhs kann hierfür als anschauliches Beispiel dienen: obgleich täglich ausgeführt, ist es nur bedingt möglich, die einzelnen Schritte des Zubindens zu verbalisieren. Beim prozeduralen Lernen handelt es sich insofern um eine Art des impliziten Lernens¹¹ (Anderson, 2013, S. 159; für eine Darstellung der unterschiedlichen Gedächtnisarten vgl. Abbildung 2).

¹¹ Implizites erlerntes Wissen ist keinem direkten Gedächtnisabruf zugänglich, sondern zeigt sich (indirekt) bei der Bearbeitung von Aufgaben (Anderson, 2013, S. 156ff.); zur Frage nach der mentalen Integration impliziten und expliziten Wissens bei Lehrkräften vgl. Stern, 2009.

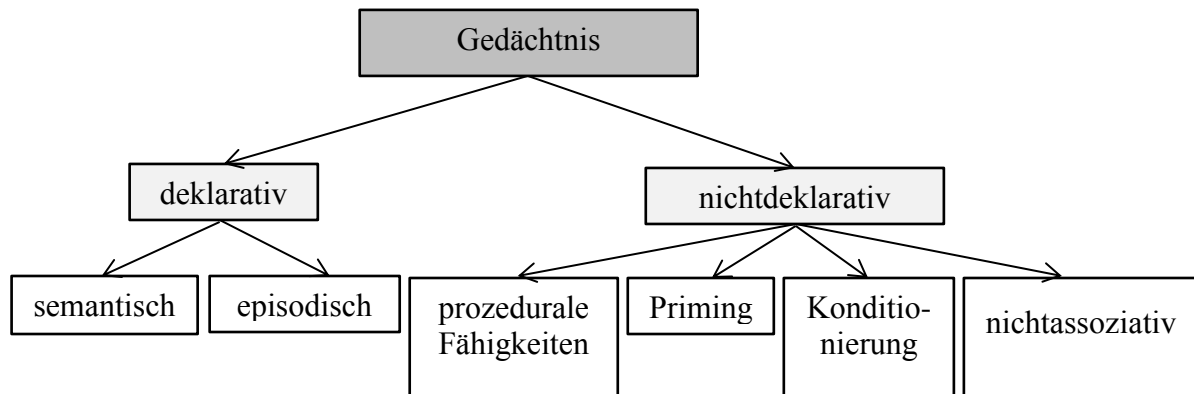


Abbildung 2: Arten des Gedächtnisses (eigene Darstellung in Anlehnung an Squire, 1987 zitiert nach Anderson, 2013, S. 160)

Zur strukturellen Unterscheidung unterschiedlicher Wissensarten wird häufig auch auf Ryle (1949/2000) rekurriert, der die Begriffspaarung „*Knowing How and Knowing That*“ (S. 28) verwendet. Die Auffassung, ‚intelligentem‘ Handeln im Sinne von Können (Bromme, 1992/2014, S. 129) gehe immer eine absichtsvolle Zurkenntnisnahme von handlungsrelevantem Wissen – „‘maxims‘, ‘imperatives‘, or ‘regulative propositions‘ as they are sometimes called“ (Ryle, 1949/2000, S. 30) – voraus, lehnt Ryle ab (ebd.). Demnach kann zwar von einem Zusammenhang dieser Wissensarten ausgegangen werden, bezogen auf das Handeln jedoch nicht im Sinne eines ‚Sich-zunächst-Vergegenwärtigens und dann Ausführens‘. Vielmehr, so Ryle weiter, wird das knowing that durch sich wiederholende Praxis in ein knowing how überführt, sodass es sich in flüssigem Handeln als *eine Operation* zeigt (ebd., S. 40f.). Es liegt nahe, das knowing that (Wissen, dass ...) und das knowing how (Wissen, wie ...) mit den Begriffen des deklarativen und des prozeduralen Wissens gleichzusetzen. Jedoch weist Bromme (1992/2014) darauf hin, dass das prozedurale Wissen „eine bestimmte Logik und Qualität des Handelns [bezeichnet], die man von außen rekonstruieren kann, die aber nicht notwendig in Form von Regeln intern repräsentiert ist“ (S. 129) und daher der Wissensbegriff „in einem anderen Sinn gebraucht wird als in den Hauptströmungen der gegenwärtigen Kognitionspsychologie“ (ebd., S. 130), was in der Expertiseforschung zu berücksichtigen sei (ebd.). Auf diese Überlegung wird im weiteren Verlauf des Kapitels noch einmal zurückgegriffen.

Es stellt sich die Frage, wie solche automatischen Umsetzungen des deklarativen Wissens in ein flüssiges Handeln (Baumert & Kunter, 2006, S. 505f.; Pfeiffer, 2008, S. 77) dargestellt werden können, in welchem strukturellen Zusammenhang also deklaratives und prozedurales Wissen stehen. Hierzu wird die *ACT*-Theorie (*Adaptive Control of Thought*) von Anderson (1983) herangezogen, die systematisiert, wie Gedächtnisspuren in Anforderungssituationen aktiviert werden (Abbildung 3).

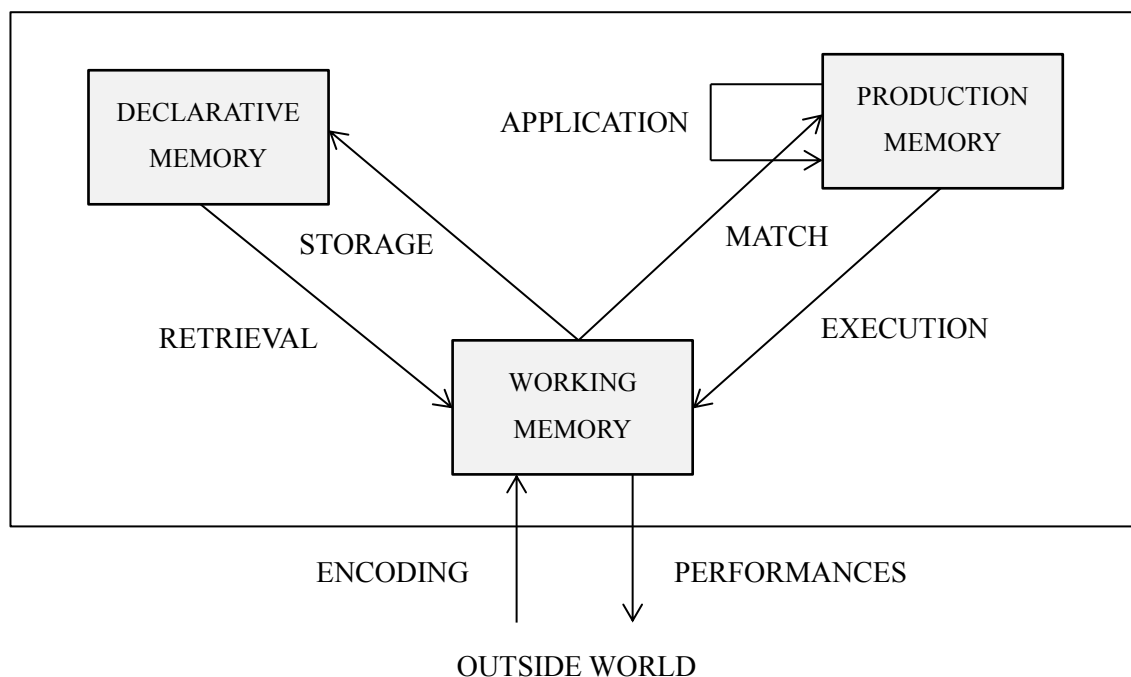


Abbildung 3: Strukturkomponenten und deren Zusammenhang in der ACT-Theorie (eigene Darstellung nach Anderson, 1983, S. 19)

In diesem Modell setzt eine Reaktion auf einen Reiz der Außenwelt die Aktivierung des deklarativen Gedächtnisses und den Abruf des notwendigen deklarativen Wissens in das Arbeitsgedächtnis einerseits sowie das Auffinden des passenden prozeduralen Wissens andererseits voraus, um ein adäquates Handeln zu ermöglichen. Die Unterscheidung in ein voneinander entkoppeltes deklaratives und prozedurales Gedächtnis ist eine zentrale Annahme der Architektur der ACT-Theorie, wodurch sich diese auch von anderen kognitionswissenschaftlichen Modellen unterscheidet, wie dem von Newell und McDermott (1975), dem eine ausschließlich prozedurale Wissensrepräsentation zugrunde liegt (Tergan, 1986, S. 153). In der

ACT-Theorie wird deklaratives Wissen in kognitiven Einheiten repräsentiert; eine solche „bedeutungsvolle Informationseinheit“ (Zimbardo & Gerrig, 2004, S. 305) wird auch als *Chunk* bezeichnet (Anderson, 2013, S. 205f.). Das prozedurale Wissen umfasst Wenn-dann-Regeln, die das deklarative Wissen in eine Produktionsstruktur überführen sollen, was Anderson (1983) mit dem Begriff der *Prozeduralisierung* so beschreibt:

„[A]ll knowledge initially comes in declarative form and must be interpreted by general procedures. However, by performing task, *proceduralization* gradually replaces the interpretive application with productions that perform the behavior directly. For example, rather than verbally rehearsing the side-angle-side rule in geometry and figuring out how it applies to a problem, a student eventually has a production that directly recognizes the application of side-angle-side.“ (S. 34)

Durch stetiges Anwenden von Handlungsregeln findet über den Vorgang der Prozeduralisierung schließlich ein Kompositionsprozess statt, der die einzelnen Produktionsschritte in einen einzigen Schritt überführt (ebd.). Dieser äußert sich in einer flüssigen Handlung und ist weiterführender Verfeinerung – „*Production tuning*“ (ebd.) – zugänglich.

Die Unterscheidung in deklaratives und prozedurales Lehrerwissen im Sinne der hier referierten kognitionspsychologischen Theorien, teilweise unter ausdrücklichem Rekurs auf die ACT-Theorie (ebd.), findet sich in unterschiedlichen Studien (vgl. u. a. Baumert & Kunter, 2011a, S. 33ff.; Krauss et al., 2008b, S. 226; Lindmeier et al., 2012, S. 105; Riese & Reinhold, 2010, S. 172; Schmelzing, 2010, S. 25ff.; Schmelzing et al., 2010, S. 191f.; Tepner et al., 2012, S. 16f.). Im Wesentlichen wird die Auffassung geteilt, dass – in einer Formulierung von Weinert et al. (1990) – „*Wissen ... nicht im umgangssprachlichen Sinne mit deklarativem Wissen gleichgesetzt werden [darf], sondern ... – gerade im Bereich des Unterrichts – zu einem erheblichen Teil als prozedurales Wissen (Routinen, Fertigkeiten, Können) angesehen werden [muss]*“ (S. 188). Jedoch liegt den Ansätzen trotz zum Teil vergleichbarer theoretischer Bezugspunkte nicht durchgängig das gleiche (begriffliche) Verständnis zu den Wirkzusammenhängen der Wissensarten zugrunde (Gramzow et al., 2013, S. 16ff.), was eine direkte Vergleichbarkeit der Ansätze erschwert. Beispielsweise verwenden Baumert und Kunter (2011a, S. 33) in der COACTIV-Studie den Begriff des Könnens als Synonym für das prozedurale Wissen, das „Fertigkeiten, Fähigkeiten, Können und Handlungsrouninen zu allen Aspekten des Unterrichtens [umfasst] (vgl. Weinert, Schrader & Helmke, 1990)“ (Krauss et al., 2008b, S. 226). Hierzu wird die Ausdifferenzierung des Wissens in theore-

tisch-formales und praktisches Wissen nach Fenstermacher (1994) herangezogen (Baumert & Kunter, 2011a, S. 35), um letzteres als ein erfahrungsbasiertes und kontextspezifisches Wissen (*knowledge in action*) zu charakterisieren, das als implizites Wissen dennoch „im praktischen Diskurs durch die Urteilskraft des professionell Lehrenden rechtfertigungsfähig [bleibt] (Hiebert, Gallimore & Stigler, 2002; Munby, Russel & Martin, 2001; Neuweg, 2005, S. 215 f.)“ (ebd.). Baumert und Kunter (2006, S. 483) gehen daher davon aus, dass ein solches praktisches Wissen teilweise mental propositional repräsentiert und insofern selektiv auch einem bewussten Abruf zugänglich ist, was in Ansätzen zur nichtpropositionalen Form der Repräsentation angezweifelt wird, nach der ein implizites Wissen vom Handelnden nicht beschrieben werden kann (Riese & Reinhold, 2010, S. 172; vgl. hierzu beispielsweise den Ansatz des *tacit knowing view* bei Neuweg, 2002). Bei Tepner et al. (2012) wird das prozedurale Wissen als die „Kenntnis von (Unterrichts-)Handlungen und -prozessen als dem *Wissen, wie etwas abläuft*“ (S. 17) beschrieben und die Situationsspezifität dieses Wissens betont (ebd.); Näheres zur Handlungswirksamkeit wird nicht ausgeführt. Riese und Reinhold (2010) bezeichnen prozedurales Wissen als „handlungsorientiertes Wissen“ (S. 172) und nehmen an, dass ein solches Wissen tentativ als Indikator für das Handeln von (hier: Physik-)Lehrkräften angesehen werden kann (ebd., S. 182).

Schon die Darstellung dieser ausgewählten Vorstellungen zum prozeduralen Wissen (vgl. auch Kapitel 2.2.3 für eine ausführlichere Gegenüberstellung unterschiedlicher Positionen mit spezifischem Bezug auf das fachdidaktische Wissen) lässt zwar im Kern vergleichbare, aber in Teilen doch unterschiedliche Positionen vor allem zur Handlungswirksamkeit des Professionswissens erkennen. Für eine systematische Klärung der Wirkzusammenhänge des Wissens, Könnens und Handelns von Lehrkräften bietet sich die Darstellung der Konzepte des Lehrerwissens bei Neuweg (2014, S. 585) an (Abbildung 4).

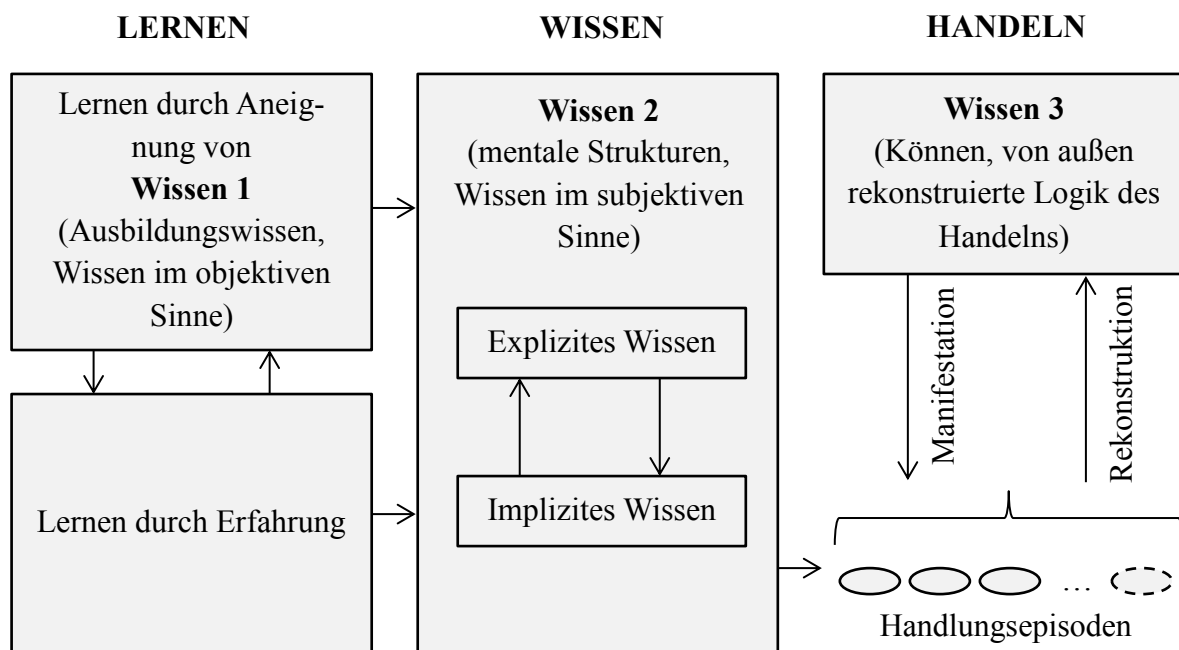


Abbildung 4: Konzepte des Lehrerwissens (eigene Darstellung in Anlehnung an Neuweg, 2014, S. 585)

In diesem Modell bezeichnet Wissen 1 „häufig das kodifizierte, mehr oder weniger systematische und insbesondere in der Ausbildung anzueignende *Professionswissen* von Lehrern“ (ebd., S. 584). Wissen 2 ist das von den Lehrpersonen mental repräsentierte Wissen; dieses entspricht nicht notwendigerweise dem objektiven Wissen (ebd.). Unter Rekurs auf Ryle (1949/2000) und Bromme (1992/2014) wird das Wissen 3 schließlich als Lehrerwissen beschrieben, bei dem es sich „in diesem Sinne ... aber nicht um das Wissen des Lehrers [handelt], sondern um das Wissen des Forschers, der die Logik des *Handelns* (!) von außen rekonstruiert“ (Neuweg, 2014, S. 585).

Von besonderem Interesse für die Frage nach dem Zusammenhang zwischen Wissen und Können ist in dieser Modelldarstellung das Wissen 2 als „*Wissen im subjektiven Sinne*“ (ebd., S. 584), da es „die eigentliche Handlungsressource [ist]“ (Vogelsang & Reinhold, 2013, S. 107). Unter Rückbezug auf die ACT-Theorie wird deutlich, dass das explizite Wissen 1 dem deklarativen Wissen bei Anderson (1983, S. 23f.) entspricht (Vogelsang & Reinhold, 2013, S. 110), also dem Wissen, „dass [*sic*] bewusst abgerufen, verbalisiert und zu Regeln geformt werden kann“ (ebd.) und das über Prozeduralisierung in ein Wissen weiter-

entwickelt wird, „dass [*sic*] die tatsächliche Ressource für Handeln darstellt“ (ebd.). Im Sinne der Unterscheidung von hypothetischen Positionen, ob und wie ein solches explizites Wissen auf das eigentliche Können von Lehrpersonen wirkt, nimmt „[d]iese Darstellung ... also eine *Transformationsposition* zum Zusammenhang von explizitem Wissen und Können ein“ (ebd.).¹² Vogelsang und Reinhold problematisieren hierbei:

„Eine Operationalisierung anhand einer theoretischen Wissenstopologie birgt allerdings die Gefahr, dass zum einen bestimmte Wissensdimensionen als Ressourcen angenommen werden, obwohl sie es nicht sind, und zum anderen Wissen erfasst wird, das nicht als Produktionsregel für das Unterrichten wirksam werden kann“ (Ebd., S. 111)

Wird davon ausgegangen, dass solche Produktionsregeln „z. B. dann erkennbar [werden], wenn Items so klassifiziert werden, dass sie sich auf Wissensmerkmale – wie konditional, deklarativ oder prozedural (Tepner et al., 2012) – beziehen, die bestimmte Phasen des Handelns ... adressieren sollen“ (ebd.), enthält dies weitreichende Implikationen sowohl für die Modellierung als auch für die Operationalisierung des professionellen Wissens und Könnens von Lehrkräften. So kann zwar davon ausgegangen werden, dass Wissen 2 nach Neuweg (2014) „in abstrakter Form durch Befragung erfasst oder aber durch situierte Aufgaben getestet werden [kann]“ (S. 587), jedoch muss die Handlungsvalidität der operationalisierten Konstrukte kritisch diskutiert werden (Vogelsang & Reinhold, 2013, S. 115f.). Wie Neuweg (2014) betont, dürfte daher „[d]ie Lehrerwissensforschung ... umso ertragreicher werden, je näher sie an die beruflichen Anforderungen heranrückt“ (S. 604).¹³

Zur Taxonomie des Lehrerwissens lässt sich zusammenfassen, dass das Professionswissen von Lehrkräften im Anschluss an eine wissensbasierte Expertiseforschung als Expertenwissen konzeptualisiert wird, „über das nur Angehörige einer bestimmten Profession verfügen“ (Vogelsang & Reinhold, 2013, S. 105; vgl. auch Besser & Krauss, 2009, S. 75). Hierzu wird auf die Kategorien von Shulman (1986, 1987) zurückgegriffen, wobei in der deutschsprachi-

¹² Der Gegensatz zur Transformationsposition stellt „eine *Position des strukturellen Unterschieds* [dar], nach der das Vorhandensein von Können auf einer anderen kognitiven Ressource beruht, [*sic*] als das Vorhandensein von in einem Test erfassbaren, expliziten Wissen“ (Vogelsang & Reinhold, 2013, S. 108).

¹³ Umgekehrt kann bei einem handlungsnah erfassten Wissen 3 nach Neuweg (2014, S. 584f.) problematisiert werden, dass „Rückschlüsse auf die Wirksamkeit der Ausbildung und Schlussfolgerungen für ihre Gestaltung ... aber wiederum schwieriger [werden], weil unklar ist, wo und wie die betreffende Person das gemessene ‚Wissen‘ (eigentlich: Können) angeeignet hat“ (S. 587). Für eine tiefer gehende kognitionstheoretische Reflexion zu solchen „Bruchstellen zwischen Wissen und Können“ (Neuweg, 2002, S. 10) vgl. ebd.

gen Forschungstradition vor allem die hervorgehobene Bedeutung des Fachwissens, des fachdidaktischen Wissens sowie des pädagogischen Wissen betont wird (Besser & Krauss, 2009, S. 78f.; Weißeno et al., 2013, S. 190). Es wird angenommen, dass das Professionswissen als kognitive Ressource für das professionelle Unterrichtshandeln dient (Vogelsang & Reinhold, 2013, S. 105). Entlang eines breiten Wissensverständnisses wird das Lehrerwissen hierzu unter Rückbezug auf Anderson (1983) häufig in deklaratives und prozedurales Wissen unterschieden (Gramzow et al., 2013, S. 16), obgleich begriffliche Unschärfen und theoretische ‚Brüche‘ diskutiert werden (Neuweg, 2002, 2014). Mit Verweis auf das Modell von Neuweg (2014, S. 585) schlussfolgern Vogelsang und Reinhold (2013) daher auch, „dass sich die Frage nach dem Verhältnis von theoretischem Wissen I [Ausbildungswissen] und Können nicht eindeutig klären lässt“ (S. 115). Gleichzeitig, so König (2010), hat sich „in gewisser Hinsicht in der Diskussion die pragmatische Einteilung in deklaratives und prozedurales Wissen durchgesetzt (Bromme 2001; Klieme et al. 2003)“ (S. 62). An einer solchen Aufteilung wird daher auch in der vorliegenden Arbeit festgehalten. Es wird davon ausgegangen, dass eine Unterscheidung der beiden Wissensarten bereits bei der Modellierung berücksichtigt werden sollte (ebd., S. 63; vgl. auch Kapitel 3.2.3), was in der Folge wiederum Implikationen für die Operationalisierung und Instrumentenentwicklung mit sich bringt. Denn während erwartet wird, dass deklaratives fachliches und fachdidaktisches Wissen durch schriftliche Tests adäquat erfasst werden kann (Vogelsang & Reinhold, 2013, S. 107), müssen für die Erhebung handlungsnaher Kognitionen andere forschungsmethodische Zugänge angedacht werden. Hierauf ist im weiteren Verlauf der Arbeit besonders einzugehen (vgl. auch Kapitel 2.3.3).

2.2 Das politikdidaktische Wissen von Lehrkräften

Eine zunächst generische Betrachtung des professionellen Wissens von Lehrkräften wie im vorigen Kapitel hat Grenzen, da „[d]ie Logik der verhandelten Sache ... sich nicht ohne Weiteres aus fachunabhängigen Sichtstrukturen des Unterrichts [erschließt]“ (Baumert et al., 2011, S. 13) und „Fächer nicht beliebige Wissenskonglomerate darstellen, sondern sachlogische Systeme, die Schüler aktiv und konstruktiv erwerben müssen“ (Weinert, 2001, S. 27). Auch die Lehrkraft benötigt fachbezogenes Wissen, um die Schüler/-innen beim Wissenserwerb unterstützen zu können, womit die Notwendigkeit einer domänenspezifischen Betrachtung der Kognitionen angesprochen ist, die Shulman (1986) wie folgt erläutert:

„In their necessary simplification of the complexities of classroom teaching, investigators ignored one central aspect of classroom life: the subject matter.

This omission also characterized most other research paradigms in the study of teaching. Occasionally subject matter entered into the research as a context variable—a control characteristic for subdividing data sets by content categories (e.g., ‘When teaching 5th grade mathematics, the following teacher behaviors were correlated with outcomes. When teaching 5th grade reading, ...’). But no one focused on the subject matter content itself.

Why this sharp distinction between content and pedagogical process?“ (S. 6)

Forschungsprogramme wie beispielsweise PKP (Oberle et al., 2012) und COACTIV (Kunter et al., 2011) betonen daher grundsätzlich die Domänenspezifität der Modellierungen, die „im Einklang mit Forderungen nach einer stärkeren Untersuchung *fachspezifischer* Aspekte in der Unterrichtsforschung [steht], die selbst von pädagogisch-psychologischer Seite erhoben werden (Helmke, 2003; Mayer, 2004)“ (Krauss et al., 2008b, S. 227). So weist Weschenfelder (2014a) für die PKP-Studie darauf hin, dass es „[b]esonders hinsichtlich der Bedeutung fachlicher Inhalte und der Gestaltung kognitiv aktivierender und gehaltvoller Lernumgebungen ... fachdidaktischen Wissens [bedarf], für das ein fachwissenschaftliches Wissensfundament notwendig scheint“ (S. 70). Lehrer/-innen müssen daher über „ein besonderes unterrichts- und schülerbezogenes fachliches Wissen“ (Baumert & Kunter, 2011a, S. 37) verfügen, um solche Lernumgebungen gestalten zu können, da „[d]as Fachwissen des Lehrers ... nur vermittelt über seine Handlungen im Unterricht auf den Lernerfolg der Schüler wirken [kann]“ (Neuweg, 2014, S. 590). Dieses spezielle Wissen, das fachliche Kenntnisse und pädagogisches Wissen inkorporiert, wird in der deutschsprachigen Forschungstradition fachdidaktisches Wissen genannt (ebd.) und in seiner besonderen Stellung innerhalb der Taxonomie des Professionswissens hervorgehoben (vgl. auch Baumert & Kunter, 2006, S. 496; Gramzow et al., 2013, S. 8; Krauss et al., 2008b, S. 227; Schmelzing et al., 2010, S. 191; Shulman, 1987, S. 8; Tepner et al., 2012, S. 13). Es ist das Wissen darüber, *wie* etwas gelehrt werden kann (Weinert et al., 1990, S. 177).

Konzeptualisierungen des fachdidaktischen Wissens in der deutschsprachigen Forschung lassen sich dabei meist einem *transformativen Modell* (Gess-Newsome, 1999) zuordnen, da sie annehmen, dass das fachdidaktische Wissen aus der Transformation des pädagogisch-

psychologischen Wissens und des Fachwissens entsteht (Gramzow et al., 2013, S. 10).¹⁴ Demgegenüber ist „das fachdidaktische Wissen im Integrativen Modell als Kreuzung von Fachwissen, pädagogischem Wissen und kontextuellem Wissen dargestellt ..., [es wird] selbst jedoch nicht als eigenständige Dimension des Professionswissens beschrieben“ (Tepner et al., 2012, S. 13). Obwohl „in der Darstellung des transformativen Modells von Gess-Newsome (1999) allein PCK das für die Unterrichtspraxis relevante Wissen beinhaltet“ (Gramzow et al., 2013, S. 10), lassen sich nicht alle Konzeptualisierungen, die in der deutschsprachigen Literatur zu finden sind, unmittelbar einem transformativen Modell zuordnen, da häufig auch die weiteren Wissensdimensionen als ebenso bedeutsam für das Unterrichtshandeln angesehen werden (ebd.). So wird in COACTIV unter anderem das Wissen über „Inszenierungsmuster von Unterricht“ (Baumert & Kunter, 2011a, S. 39) oder auch das Wissen über „Unterrichtsmethoden im weiten Sinne“ (ebd.) den Facetten des pädagogischen Wissens zugeordnet (ebd.).

Die unterschiedliche Verwendung von theoretischen Rahmenmodellen (integrativ versus transformativ) zum Professionswissen von Lehrkräften lässt zum einen vermuten, dass – wie es Gramzow et al. (2013) für das mathematisch-naturwissenschaftliche Wissen von Lehrkräften formulieren – „in den verschiedenen Studien für die MINT-Fächer national wie auch international keine einheitliche Modellierung dieses Wissensbereichs zugrunde liegt“ (S. 8; für eine Übersicht über Konzeptualisierungen im englischsprachigen Raum vgl. Park & Oliver, 2008). Zum anderen deuten sich in den Konzeptualisierungen eines transformativen Modells ‚unscharfe Ränder‘ zwischen den Wissenskategorien nach Shulman (1986, 1987) an, die Neuweg (2014) am Beispiel der „Brüchigkeit der Unterscheidung“ (S. 591) zwischen dem Fachwissen und dem fachdidaktischen Wissen so beschreibt:

„Jeder Fachwissenstest nötigt die Versuchsperson zur Repräsentation des eigenen Wissens und wird daher auch vor dem Hintergrund konkreter Vorstellungen davon entwickelt, was als qualitätsvolle Repräsentation von Fachwissen gelten kann. Die Fähigkeit, das eigene Wissen darzustellen – ein Kernbestandteil fachdidaktischen Wissens –, lässt sich theoretisch wie praktisch also kaum vom ‚reinen‘ Fachwissen unterscheiden (Borko & Putnam, 1996, S. 691f.). Und umgekehrt ist die Fähigkeit, Wissen vermitteln zu können, ein zentraler Indika-

¹⁴ Diese Unterscheidung enthält Implikationen für die Erwerbs- und Vermittlungsprozesse des Professionswissens, weil nach dieser Modellvorstellung davon ausgegangen wird, dass das fachdidaktische Wissen „eine eigenständige Vermittlung im Rahmen der Lehrerbildung ein[fordert]“ (Schmelzing et al., 2010, S. 191). Für eine Diskussion, inwieweit die fachbezogenen Wissensbestandteile tatsächlich auch wirksam in der Lehrerbildung an Universitäten vermittelt werden können, vgl. Hascher (2014).

tor für die Tiefe des Fachwissens, wie jeder Lehrer weiß, dem erst im Unterricht vollends deutlich wird, ob und wie gut er sein Thema beherrscht.“ (Ebd.)

Der Umstand, dass also das fachdidaktische Wissen zwar als zentraler Bestandteil des Professionswissens angesehen wird, das Fachwissen jedoch Voraussetzung – im Sinne einer Grundlage und nicht lediglich als korrelierendes Konstrukt – hierfür ist (Baumert et al., 2011, S. 15; Krauss et al., 2008b, S. 226f.), stellt eine Herausforderung sowohl für die theoretische Konzeptualisierung als auch für die Messung dar. Ein Beispiel-Item (in der späteren Testdurchführung nicht verwendet) aus dem PKP-Forschungsprogramm (Oberle et al., 2012) kann dies verdeutlichen (Abbildung 5).

Die Schüler/-innen diskutieren im Unterricht kritisch über die EU.

Chantal sagt:

Die Wahlen zum Europäischen Parlament sind belanglos, weil das Europäische Parlament nichts zu sagen hat.

Welche Fehlvorstellung hat Chantal?

- ☐ Sie weiß nicht, dass das Europäische Parlament aus Abgesandten der nationalen Parlamente besteht.
- ☒ Sie hat nicht realisiert, dass das Europäische Parlament bei der Gesetzgebung der EU meistens das Vetorecht hat.
- ☐ Sie weiß nicht, dass der Sinn von Wahlen darin besteht, das Interesse an Politik zu steigern.

Abbildung 5: Beispiel-Item zur Erfassung des lernprozessbezogenen politikdidaktischen Wissens in der PKP-Studie (eigene Darstellung nach Weschenfelder, 2014a, S. 171)

Das Beispiel-Item verdeutlicht, dass das Vorhandensein fachwissenschaftlicher Kenntnisse eine unabdingbare Voraussetzung dafür ist, die Frage korrekt beantworten zu können. Gleichzeitig lässt sich vermuten, dass „mit der fachdidaktischen Kompetenz im Grunde immer auch zugleich die Tiefe des Verständnisses des Fachwissens gemessen [wird]“ (Neuweg, 2014, S. 592). Empirische Befunde unterschiedlicher Forschungsprojekte spiegeln dieses durchaus uneinheitliche Bild zum Zusammenhang zwischen Fachwissen und fachdidaktischem Wissen wider. So konnten Riese und Reinhold (2010, S. 182ff.) für das physikdidaktische Wissen sowie Krauss et al. (2011, S. 148) für das mathematikdidaktische Wissen „zeigen, dass ihre Operationalisierung des o. g. Bereichs als Fachdidaktisches

Wissen neben Fachwissen und Pädagogischem Wissen ein eigenständiges Konstrukt innerhalb des Professionswissens darstellt“ (Gramzow et al., 2013, S. 8). Gleichzeitig wird hinsichtlich der von Krauss et al. (2011, S. 149) für die COACTIV-Studie berichteten substanziellen Korrelation von .63 zwischen dem Fachwissen und dem mathematikdidaktischen Wissen im Zusammenhang mit teilweise höheren Korrelationen zwischen dem Fachwissen und den Subskalen des fachdidaktischen Wissens (ebd., S. 150; vgl. auch Krauss, Baumert & Blum, 2008a) diskutiert, „ob dadurch die Unterscheidung zwischen Fachwissen und fachdidaktischer Kompetenz eher gestützt oder aber in Frage gestellt wird“ (Neuweg, 2014, S. 592).

In anderen Studien wird das Fachwissen und das fachdidaktische Wissen daher zu einem übergreifenden Konstrukt enger zusammengefasst. So bilden Hill, Schilling und Loewenberg Ball (2004) Facetten des mathematischen Fachwissen und des mathematikdidaktischen Wissens von Grundschullehrkräften als gemeinsames „mathematical knowledge for teaching“ (S. 12, vgl. auch Kleickmann et al., 2013, S. 91) ab.¹⁵ Lindmeier et al. (2012, S. 101f.) führen zu den Aufgaben zum fachdidaktischen lernprozessbezogenen Wissen im *TEDS-M*-Projekt (Blömeke et al., 2010a; Blömeke et al., 2010b) an, dass deren korrekte Beantwortung kein fachdidaktisches Wissen erfordere und schlussfolgern für ihre Modellierung: „Da schulnahes Fachwissen und fachdidaktisches Wissen in bisherigen Studien empirisch schwer zu trennen war (Hill et al. 2004; Krauss et al. 2008a), werden diese beiden Aspekte in dem vorliegenden Modell zu einer einzigen Komponente Basiswissen zusammengefasst (basic knowledge, BK)“ (Lindmeier et al., 2012, S. 105). Gleichzeitig betonen sie, dass weiterhin eine Kompatibilität zur Zweiteilung Shulmans (1986, 1987) begründbar sei, da das Basiswissen „konsequent als Wissenskomponente verstanden und das dazugehörige Faktenwissen und prozedurale Wissen ... so gemessen werden [kann], wie es in bisherigen Projekten geschehen ist (Hill et al. 2004; Krauss et al. 2008b; Blömeke et al. 2008)“ (Lindmeier et al., 2012, S. 105).

Hinsichtlich der Frage nach der internen Strukturierung des fachdidaktischen Wissens lässt sich zusammenfassen, dass eine Vielzahl deutschsprachiger Ansätze zum Professionswissen im Sinne der Taxonomie Shulmans (1986, 1987) an der theoretischen Abgrenzung zum

¹⁵ Krauss et al. (2008a) betonen dazu jedoch, dieses sei „distinguished from *subject matter knowledge* or ‘pure’ content knowledge“ (S. 877).

Fachwissen festhält, sich diese Zweiteilung jedoch nicht in allen Fällen durch empirische Befunde bestätigen lässt (Neuweg, 2014, S. 592). Auch theoretisch scheint der Zusammenhang zwischen Fachwissen und fachdidaktischem Wissen nicht eindeutig geklärt zu sein. Im Anschluss an die Konzeptualisierung des politikdidaktischen Wissens in der PKP-Studie (Oberle et al., 2012) wird in der vorliegenden Arbeit ebenfalls davon ausgegangen, „dass fachwissenschaftliches und fachdidaktisches Wissen zwei getrennte Komponenten des Wissenskonstruktes darstellen, was für eine differenzierte Modellierung und Analyse bedeutsam ist“ (Weschenfelder, 2014a, S. 70). Gleichzeitig ist hierauf insbesondere in der Interpretation der Zusammenhangsmaße der Skalen zum Fachwissen und zum fachdidaktischen Wissen einzugehen (vgl. auch Kapitel 7.3). Durch die inhaltliche Nähe des Fachwissens zum fachdidaktischen Wissen ergibt sich die Frage, wie letzteres schließlich konzeptualisiert werden kann, was im Folgenden in den drei Strukturdimensionen *Facetten* (Kapitel 2.2.1), *Inhaltsbereiche* (Kapitel 2.2.2) und *Wissensarten* (Kapitel 2.2.3) näher analysiert werden soll.

2.2.1 Strukturdimension *Facetten*

Die spezifische Form des fachdidaktischen Wissens als Amalgam aus Fachwissen und pädagogischem Wissen (Shulman, 1987, S. 8) beschreibt Shulman in einem Artikel von 1986:

„Within the category of pedagogical content knowledge I include, for the most regularly taught topics in one’s subject area, the most useful forms of representation of those ideas, the most powerful analogies, illustrations, examples, explanations, and demonstrations—in a word, the ways of representing and formulating the subject that make it comprehensible to others.

Pedagogical content knowledge also includes an understanding of what makes the learning of specific topics easy or difficult: the conceptions and preconceptions that students of different ages and backgrounds bring with them to the learning of those most frequently taught topics and lessons. If those preconceptions are misconceptions, which they so often are, teachers need knowledge of the strategies most likely to be fruitful in reorganizing the understanding of learners, because those learners are unlikely to appear before them as blank slates.“ (S. 9f.)

Es werden in diesem Zitat vor allem zwei Facetten des fachdidaktischen Wissens betont, nämlich zum einen das Wissen über adäquate Darstellungs- und Vermittlungsstrategien und zum anderen das Wissen über die domänenspezifischen (Fehl-)Vorstellungen der

Schüler/-innen (Tepner et al., 2012, S. 14). Ein weiterer von Shulman (1986, 1987) eingebrachter Aspekt fachdidaktischen Wissens, das in Kapitel 2.1.1 bereits genannte curriculare Wissen, wird in der Forschung nur selten aufgegriffen (vgl. auch Tepner et al., 2012, S. 8f.). In den Übersichten von Park und Oliver (2008, S. 265) über die konzeptualisierten Facetten des fachdidaktischen Wissens in unterschiedlichen Studien werden die beiden erstgenannten Facetten in ihrer zentralen Stellung hervorgehoben und auch die aktuellere Zusammenschau von Schmelzing (2010) bestätigt dies (Tabelle 1).

Tabelle 1: Übersicht über Facetten fachdidaktischen Wissens in unterschiedlichen Studien (eigene Darstellung nach Schmelzing, 2010, S. 23)

Referenz bzw. Facette	Instruktions- und Vermittlungsstrategien	Schüler und Schülerkognitionen	Ziel von gelehrtem Fachinhalt	Lehrplan	Bewertung und Beurteilung	Medien	Fachwissen	Kontext	Pädagogik
Shulman (1986)	x	x							
Tamir (1988)	x	x		x	x				
Smith & Neale (1989)	x	x	x						
Grossman (1990)	x	x	x	x					
Marks (1990)	x	x				x	x		
Cochran et al. (1993)		x					x	x	x
Geddis et al. (1993)	x	x							
Fernandez-Balboa & Stiehl (1995)	x	x	x	x			x	x	
van Driel et al. (1998)	x	x							
Magnusson et al. (1999)	x	x	x	x	x				
Hashweh (2005)	x	x	x	x	x		x	x	x
Loughran et al. (2006)	x	x	x				x	x	x
Krauss (2006)	x	x							
Park & Oliver (2008)	x	x	x	x	x				
Lee & Luft (2008)	x	x	x	x	x	x			
Rohaan (2009)	x	x	x						
Riese (2009)	x	x	x	x	x	x			
van Dijk & Kattmann (2010)	x	x							

Es gilt zu berücksichtigen, dass eine solche Übersicht zunächst lediglich Übereinstimmungen auf Ebene der Sichtstruktur aufzeigen kann. Es wird im Hinblick auf die in Kapitel 2.2 angesprochene Domänenspezifität des fachdidaktischen Wissens deutlich, dass man das, „[w]as z.B. ‚useful forms of representation‘ oder ‚typical misconceptions‘ sein könnten, ... nur fachspezifisch einschätzen und begründen können [wird]“ (Krauss et al., 2008b, S. 228). Für die vorliegende Arbeit ist daher von Interesse, was aus domänenspezifischer Sicht unter den in Tabelle 1 genannten Facetten bzw. Referenzen des fachdidaktischen Wissens verstanden wird, wie diese also inhaltlich ‚gefüllt‘ werden. Zu diesem Zweck soll gegenübergestellt werden, wie das fachdidaktische Wissen über Instruktions- und Vermittlungsstrategien sowie über Kognitionen der Schüler/-innen in unterschiedlichen Untersuchungen in anderen Domänen konzeptualisiert wird, um Gemeinsamkeiten und Unterschiede erkennen zu können, die für die weiterführenden politikdidaktischen Überlegungen besonders relevant erscheinen (Tabelle 2). Die Aufstellung enthält zur besseren Vergleichbarkeit ausgewählte Konzeptualisierungsansätze aus der mathematisch-naturwissenschaftlichen Domäne, die sich grundständig auf die Shulman’sche (1986, 1987) Taxonomie des Lehrwissens beziehen und hierbei das fachdidaktische Wissen dergestalt beschreiben, dass es einer konkreten Operationalisierung (Gramzow et al., 2013, S. 9) zugänglich gemacht werden kann.

Tabelle 2: Operationalisierungen des fachdidaktischen Wissens über Instruktions- und Vermittlungsstrategien sowie über Kognitionen der Schüler/-innen in der mathematisch-naturwissenschaftlichen Domäne in unterschiedlichen Studien

Facette	Instruktions- und Vermittlungsstrategien	Kognitionen der Schüler/-innen
Tepner et al. (2012)	Umgang mit Experimenten und Modellen ¹⁶	Vorstellung der Schüler/-innen (korrekte und inkorrekte Präkonzepte) und Fehlkonzepte
Schmelzing et al. (2010)	Illustrationen, Repräsentationen und fachspezifische Instruktionsstrategien	Kognitionen der Schüler/-innen
Riese & Reinhold (2010)	Wissen über den Einsatz von Experimenten; Gestaltung und Planung von Lernprozessen	Wissen über (allgemeine) Aspekte physikalischer Lernprozesse; Beurteilung, Analyse und Reflexion von Lernprozessen
van Dijk & Kattmann (2010)	Kenntnisse von Instruktionsformen (Beispiele, Experimente, Ereignisse, Aufgaben und Situationen)	Kenntnisse der Vorstellungen und Verständnisprobleme von Schülerinnen und Schülern
Krauss et al. (2011)	Zugänglichmachen: Wissen über Erklären und Repräsentieren	Schülerinnen und Schüler: Wissen über typische Fehler und Schwierigkeiten
Brovelli, Bölsterli, Rehm & Wilhelm (2013)	lernendengerechte Instruktionsstrategien; fachgerechte Instruktionsstrategien; Umgang mit originalen Begegnungen und Experimenten	Umgang mit Vorstellungen der Schüler/-innen; kognitive Aktivierung im Hinblick auf die Veränderung von Vorstellungen
Krauss et al. (2008b)	Wissen über das Verständlichmachen von mathematischen Inhalten	Wissen über mathematikbezogene Kognitionen der Schüler/-innen

Fortsetzung

¹⁶ Die Kenntnis fachunspezifischer Methoden sowie das Wissen über die Möglichkeiten der Passung dieser Methoden und des Unterrichtsinhalts sind bei Tepner et al. (2012, S. 11) als Bestandteil des pädagogischen Wissens aufgeführt.

Facette	Instruktions- und Vermittlungsstrategien	Kognitionen der Schüler/-innen
Gramzow et al. (2013)	Wissen über verschiedene Formen von physik- und themenspezifischen Instruktions- und Vermittlungsstrategien; Wissen über deren Anwendung und Bewertung; Kenntnisse von Verständnisschwierigkeiten bei verschiedenen Formen der Inhaltsdarstellung; Kennen von geeigneten Beispielen; Wissen über die Nutzung geeigneter Arbeitsformen; Wissen über die Möglichkeiten kumulativen Wissensaufbaus und inhaltlicher Vernetzung; Wissen über Möglichkeiten kognitiver Aktivierung; Wissen über die Reflexion von Lehr-Lern-Angeboten; Wissen über Experimente und fachdidaktische Konzepte ¹⁷	Wissen über typische themenspezifische und übergreifende Vorstellungen der Schüler/-innen; Kenntnisse zur Prognose von typischen Antworten und zur Diagnose von Fehlvorstellungen; Wissen über Vorstellungen der Schüler/-innen bei der Unterrichtsplanung; Kenntnisse über die für die Lernenden üblichen schwierigen Teilgebiete; Wissen über ihre Interessensbereiche ¹⁸

Anmerkung: Soweit möglich werden die von den Autoren verwendeten Begriffe wiedergegeben, anderenfalls werden sie inhaltsgleich reformuliert. Differenzierungen der Autoren hinsichtlich der Wissenarten (deklarativ, konditional, prozedural; kognitiv – nicht kognitiv; Wissen – Können) werden für diese Zusammenschau zunächst vernachlässigt.

Es lassen sich in Tabelle 2 deutliche Ähnlichkeiten im Verständnis der Facetten des fachdidaktischen Wissens über Instruktions- und Vermittlungsstrategien sowie über Kognitionen der Schüler/-innen zur der Konzeptualisierung Shulmans (1986, 1987) erkennen, allerdings wird dieses Verständnis auf zum Teil sehr unterschiedlichem Niveau konkretisiert. Wie Gramzow et al. (2013, S. 16) feststellen, ist der direkte Vergleich daher schwierig. Während einige Konzeptualisierungen (Krauss et al., 2011; Schmelzing et al., 2010) im Wesentlichen die Auffassungen Shulmans (1986) umformulieren, deren fachspezifische Ausdifferenzierung zum Wissen über Instruktions- und Vermittlungsstrategien jedoch nicht weiter erläu-

¹⁷ In der Konzeptualisierung von Gramzow et al. (2013, S. 23f.) ist sowohl das Wissen über Experimente als auch das Wissen über fachdidaktische Konzepte als separate Facette modelliert. Für diese Übersicht wird beides jedoch dem Wissen über fachspezifische Instruktions- und Vermittlungsstrategien zugeordnet.

¹⁸ Die Wissensfacette *Kontext und Interesse* ist bei Gramzow et al. (2013, S. 24) eigenständig und wird daher zum Beispiel vom Wissen über typische Schülerfehler unterschieden. Entsprechend der Einteilung der Tabelle wird diese Wissensfacette jedoch unter der Überschrift *Schülerkognitionen* subsumiert.

tern, explizieren dies andere Untersuchungen (Riese & Reinhold, 2010, S. 172f.; Tepner et al., 2012, S. 15f.). Diese Unterschiede können mit dem bereits angesprochenen heuristischen Charakter der Taxonomie von Shulman (1986, 1987) zusammenhängen. Shulman betont zwar die Notwendigkeit einer Klärung dessen, was adäquate Instruktions- und Vermittlungsstrategien in der jeweiligen Domäne sind (Krauss et al., 2008b, S. 228), warnt aber gleichzeitig: „The great danger occurs, however, when a general teaching principle is distorted into prescription, when a maxim becomes mandate“ (Shulman, 1987, S. 11). In diesem Fall gäbe es „no single most powerful forms of representation“ (Shulman, 1986, S. 9). Die Formulierung Shulmans – „the ways of representing and formulating the subject that make it comprehensible to others“ (ebd.) – legt daher die Lesart nahe, dass Lehrpersonen über differenzierte Mittel zum Erklären und Verständlichmachen verfügen müssen, was expositorische Vermittlungsstile genauso einschließt wie Ansätze entdeckenden Lernens (Hameyer & Röber, 2011). Denn wie Helmke (2014) feststellt, ist „der erfolgreichste Unterricht selten derjenige mit einem *Maximum* an Methodenvariationen, sondern der mit einem *Optimum*“ (S. 270). Welche Methoden dies dabei in den jeweiligen Fächern sind, lässt sich teilweise in Tabelle 2 ablesen. In den naturwissenschaftlichen Konzeptualisierungen dominiert das Wissen über Experimente und Modelle zum Erklären und Repräsentieren von Unterrichtsinhalten, was sich mit empirischen Befunden der Domäne zur methodischen Wirksamkeit deckt. So zeigt die TIMSS/III-Studie (*Third International Mathematics and Science Study*) für den Physikunterricht, dass für diesen „[d]as Zusammenspiel zwischen lehrergeleiteter Konzeptentwicklung und dem Demonstrationsexperiment ... wohl das eigentliche Fundament [ist]“ (Baumert & Köller, 2000, S. 291). Insgesamt geben 56 % der befragten Schüler/-innen an, dass sie in jeder Stunde oder in den meisten Stunden der Lehrkraft dabei zusehen, wenn diese Konzepte entwickelt oder darstellt (ebd., S. 284), und „[n]ach den Aussagen von 89 Prozent der Befragten mit Physikunterricht wird dieses didaktische Vorgehen in einigen oder in den meisten Stunden durch Demonstrationsexperimente unterstützt“ (ebd.). Aber auch dem von den Schülerinnen und Schülern selbst durchgeführten Experiment wird im Physikunterricht eine entscheidende Bedeutung für den Lernerfolg beigemessen (ebd., S. 292f.).

Unter der Facette des Wissens über Kognitionen der Schüler/-innen wird in nahezu allen aufgeführten Ansätzen mindestens das Wissen über richtige und falsche Vorstellungen gefasst. Gramzow et al. (2013) benennen als weitere Facette unter anderem das Wissen über

Kontext und das Wissen zur „interessensorientierten Unterrichtsgestaltung (vgl. Korneck et al., 2010)“ (S. 24), das entsprechend der Aufteilung in Tabelle 2 ebenfalls dem Wissen über Kognitionen zugeordnet wird; es beinhaltet insbesondere Kenntnisse über die Einbettungsmöglichkeiten von Lernangeboten in lebensweltliche Kontexte (ebd.). Dies soll – wie es Kölbach und Sumfleth (2012) für das Fach Chemie formulieren – „Lernenden helfen, durch die Anbindung an alltagsweltliche Problemstellungen ein fundiertes und strukturiertes Wissen zu erlangen, und die Anwendbarkeit und Nützlichkeit des Wissens betonen (Bennet & Holman, 2002; Bennet et al. 2007)“ (S. 254; vgl. auch Kölbach & Sumfleth, 2013). Das lernprozessbezogene Wissen über das Schülerwissen wird in einigen Ansätzen (vgl. u. a. Riese & Reinhold, 2010, S. 173) ergänzt durch das eher auf das Lernprodukt bezogene Wissen über Möglichkeiten der Reflexion und Bewertung.

Die Facette des Wissens über typische (Fehl-)Vorstellungen der Lernenden zeigt nach diesem Verständnis eine deutliche Nähe zu Aspekten der *diagnostischen Kompetenz* (Jäger, 2009) oder auch der – begrifflich enger gefassten – *diagnostischen Fähigkeit* (Brunner, Anders, Hachfeld & Krauss, 2011) von Lehrkräften, wie sie beispielsweise in COACTIV (Kunter et al., 2011) konzeptualisiert wurde. Nach einem breiten Verständnis des Diagnostikbegriffs sind Bereiche diagnostischer Tätigkeiten von Lehrkräften „in der Notengebung ebenso angesiedelt wie im Umfeld der Förderung, welche eine gezielte Diagnostik voraussetzt, oder auch bei Fragen des Übergangs in weiterführende Schulen“ (Jäger, 2009, S. 106; vgl. auch Dübbelde, 2013). In der COACTIV-Studie sind die diagnostischen Fähigkeiten von Lehrkräften zwar enger gefasst, liegen aber dennoch gewissermaßen ‚quer‘ zu Facetten des fachdidaktischen Wissens und des pädagogisch-psychologischen Wissens (Brunner et al., 2011):

„Diagnostische Fähigkeiten erfordern von Mathematiklehrkräften die Integration von diagnostischem Methodenwissen, Wissen über das Potenzial von mathematischen Aufgaben und Wissen über fachbezogene Schülerkognitionen, damit sie (a) lern- und leistungsrelevante Merkmale sowie (b) Lern- und Aufgabenanforderungen für die Schülerinnen und Schüler ihrer Klassen adäquat einschätzen können.“ (S. 219)

Die empirischen Befunde in COACTIV zeigen dabei, dass „[d]iagnostische Fähigkeiten der Lehrkräfte (aufgabenbezogener Urteilsfehler und diagnostische Sensitivität) ... einen positiven Einfluss auf die Mathematikleistung ihrer Schülerinnen und Schüler [haben]“ (ebd., S. 230) und dass diese Fähigkeiten ein mehrdimensionales Konstrukt darstellen (ebd.).

Wird der Fokus auf domänenspezifische politikdidaktische Überlegungen zur Konzeptualisierung des fachdidaktischen Wissens gelegt, scheint eine Übertragung der beiden zentralen Facetten des fachdidaktischen Wissens auf das politikdidaktische Wissen im Anschluss an die Konzeptualisierungsansätze anderer Domänen – selbstverständlich in einer domänenspezifisch adaptierten Form – möglich. So fasst Weschenfelder (2014a) für die Grundlage der Konzeptualisierung des politikdidaktischen Wissens in der PKP-Studie zusammen:

„Das politikdidaktische Wissen von Politiklehrkräften bezieht sich auf die fachbezogene Gestaltung des Unterrichts, der den Erwerb politischer Kompetenzen in geeigneten Lernprozessen initiieren und unterstützen soll. Neben der begründeten Auswahl von Zielen und Inhalten sind Kenntnisse über Lehr-Lernprozesse im Fach, u. a. über die Repräsentation und Strukturierung von Inhalten und über die Bedingungen des Kompetenzaufbaus, sowie über Lehr-Lernprozesse bedeutsam (vgl. Massing, 2007, S. 292; Weißeno et al., 2010; Bromme, 1992; Shulman, 1986; Baumert & Kunter, 2006). Auch für Politiklehrkräfte kann angenommen werden, dass sie Wissen in den verschiedenen Bereichen benötigen, um erfolgreich Unterrichten [*sic*] zu können.“ (S. 103).

Folglich werden auch in der PKP-Studie die „Fähigkeiten, Fachinhalte adäquat zu repräsentieren (*knowledge of instructional strategies*) und typische Schüler(fehl)konzeptionen (*knowledge of students' understanding*) zu erkennen“ (ebd., S. 109), als zentrale Facetten des politikdidaktischen Wissens angesehen (ebd.). Es stellt sich dabei die Frage, welche domänenspezifischen Referenzpunkte es für die Konzeptualisierung des politikdidaktischen Wissens gibt. Eine Möglichkeit besteht darin, das fachdidaktische Wissen in Anlehnung an berufsbezogene Standards zu entwickeln. So finden Gramzow et al. (2013) für die Entwicklung des fachdidaktischen Wissens von angehenden Physiklehrkräften „Hinweise für eine differenzierte Modellierung ... in normativen Setzungen wie beispielsweise den Bildungsstandards (KMK, 2008) oder dem Kerncurriculum Fachdidaktik (Gesellschaft für Fachdidaktik, 2004)“ (S. 20; vgl. auch Blömeke, 2013a, S. 8), wobei für politikdidaktische Überlegungen kritisch zu reflektieren ist, inwieweit sich solche bildungsadministrativen Rahmenvorgaben hierzu eignen. Zwar zeigen sich die im *Kerncurriculum Fachdidaktik* der *Gesellschaft für Fachdidaktik* (GFD, 2004) aufgeführten Kompetenzerwartungen grundständig anschlussfähig an die angeführte Wissenstaxonomie; bei näherer Betrachtung wird jedoch deutlich, dass es weiterhin einer Konkretisierung bedarf, was die „Kenntnis und Beurteilung beispielhafter fachdidaktischer Ansätze für die Unterstützung von Lernprozessen“ (ebd., S. 3) oder die „Fähigkeit zum (exemplarischen) Planen und Gestalten eines strukturierten Lerngangs“ (ebd., S. 5) fachspezifisch auszeichnet. Dies kann nur politikdidaktisch geklärt werden.

In der PKP-Studie wird ein Vorschlag von Batt, Massing und Sarcinelli (2007; zitiert nach Weschenfelder, 2014a) für die Konzeptualisierung des politikdidaktischen Wissens herangezogen. Dieser diene auch als Basis für die Entwicklung der *Ländergemeinsamen inhaltlichen Anforderungen für die Fachwissenschaften und die Fachdidaktiken in der Lehrerbildung* der KMK (2015), die als bildungsadministrative Rahmenvorgaben die Grundlage zur Einrichtung und Evaluierung von Lehramtsstudiengängen bilden:

- „Entwicklung politischer Erziehung und Bildung sowie politikdidaktischer Ansätze und Konzeptionen
- politische Sozialisation von Jugendlichen, Probleme und Handlungsstrategien bei der Entwicklung Demokratie fördernder Unterrichtsprozesse
- Lehr-Lern-Forschung, Kompetenz-, Einstellungs-, Urteils- und Wissenserwerb, mentale Modelle und Misconceptions bei Lernenden
- Konzepte didaktischer Analyse und Planung sowie Evaluation von Unterricht in Politik/Wirtschaft
- Methoden, Arbeitstechniken und Medien für den Unterricht in Politik/Wirtschaft“ (Batt et al., 2007; zitiert nach Weschenfelder, 2014a S. 109)

Es zeichnet sich dabei ein gewisses Problem bei der Konzeptualisierung des politikdidaktischen Wissens ab, das mit einem weiterhin zu konstatierenden Theorie- und Empiriedefizit der Politikdidaktik (Manzel, 2014; Weißeno, 2012) zusammenhängen dürfte. Da nach wie vor fachspezifische Forschungserkenntnisse unter anderem zu Wirkungsweisen von unterschiedlichen Methoden, Medien und didaktischen Prinzipien der politischen Bildung fehlen (Manzel & Neumann, 2015, S. 76f.), scheint eine empirisch fundierte ‚inhaltliche Füllung‘ der Facette des Wissens über Instruktions- und Vermittlungsstrategien – wie oben am Beispiel von Experimenten in den Naturwissenschaften angesprochen – nicht möglich. Ähnliches lässt sich für die Facette des Wissens über typische (Fehl-)Vorstellungen der Lernenden feststellen. Um zu systematisieren, welche dieser Vorstellungen im Politikunterricht zu erwarten sind, bedarf es adäquater Forschungen, die bislang in der Politikdidaktik mit einigen Ausnahmen (Götzmann, 2008; Hahn-Laudenberg, Weißeno & Eck, 2010) noch nicht vorliegen (D. Neumann, 2014, S. 56f.; Weschenfelder, 2014a, S. 169f.).¹⁹ Es wird daher darauf hingewiesen, dass – ergänzend zur theoretischen Rahmung durch den Rekurs auf etablierte taxonomische Kategorien wie die von Shulman (1986, 1987)

¹⁹ Daher muss auch hinterfragt werden, ob mit dem in Abbildung 1 dargestellten Beispiel-Item aus der PKP-Studie tatsächlich das Lehrerwissen über *typische* (Fehl-)Konzepte der Lernenden erfasst wird, da es sich bei der für die Itemkonstruktion herangezogenen Aufstellung von Fehlkonzepthen bei Weißeno, Detjen, Juchler, Massing und Richter (2010) eher um *exemplarische* Fehlkonzepthen handelt.

sowie Strukturmodelle wie COACTIV (Kunter et al., 2011) – „[d]ennoch ... die normativen fachdidaktischen Überlegungen Hinweise für die Bestimmung domänenspezifischer Kompetenzdimensionen geben [können]“ (Weschenfelder, 2014a, S. 56), was sich auch auf die Konkretisierung der Facetten politikdidaktischer Kompetenzen bezieht (ebd., S. 57).

Obgleich eine solche „Normativität der zahlreichen Entscheidungen im Zuge der Kompetenzmodellierung beispielsweise bei der Operationalisierung oder im Zuge der Kompetenzerfassung beispielsweise bei der Auswahl der Messverfahren oder angesichts der Interdependenz von Zielen, Inhalten und Methoden“ (Blömeke, 2013b, S. 34) problematisiert werden kann, wird die Notwendigkeit einer normativen Verortung der so konzeptualisierten Wissensfacetten auch in anderen Domänen an unterschiedlichen Stellen betont. So weist bereits Shulman (1986) für die Facette des Wissens über Instruktions- und Vermittlungsstrategien darauf hin, dass es unterschiedlicher Strategien bedarf, um diese – „some of which derive from research whereas others originate in the wisdom of practice“ (S. 9) – näher bestimmen zu können. Krauss et al. (2008b) betonen im Rahmen der Operationalisierung des fachdidaktischen Wissens in der COACTIV-Studie, „dass es für fachdidaktische Items oft keine eindeutig normativ richtige Antwort gibt“ (S. 239), weshalb die Items durch Experten (Fachdidaktiker/-innen, Mathematiklehrer/-innen) konstruiert wurden (ebd.). Bei Riese & Reinhold (2010) wurden Test-Items auch „unter Einbezug normativer, fachdidaktischer Leitbilder (etwa Didaktische Rekonstruktion, vgl. Kattmann, Duit, Gropengießer & Komorek, 1997) konkretisiert bzw. operationalisiert“ (S. 173).

Ein normativ geprägter Diskurs in der Politikdidaktik, der ausgewählte Ansatzpunkte für eine nähere Festlegung des domänenspezifischen fachdidaktischen Wissen aufzeigen kann, lässt sich auch im Hinblick auf das unterrichtsrelevante fachdidaktische Wissen von Politiklehrkräften ausmachen (Reinhardt, 2009, 2012a; Petrik, 2009, 2012; vgl. auch Kapitel 2.3.4 für eine kritische Reflexion dieser Überlegungen vor dem Hintergrund kompetenztheoretischer Annahmen). Wenn sich dieses Wissen nach der Auffassung Shulmans (1987) durch ein didaktisches Dreieck aus Sachgegenstand, Lernenden und Lehrkraft abbilden lässt (vgl. auch Krauss et al., 2008, S. 233), zeigt sich eine gewisse Nähe zu den angenommenen Funktionen von politikdidaktischen Prinzipien (Manzel & Neumann, 2015, S. 76). So werden in der Politikdidaktik fachdidaktische Prinzipien formuliert, „die sich methodisch artikulieren lassen ... [und] das Zusammengehen des didaktischen Dreiecks (Sache, Lerner, Lehrender) [leis-

ten]“ (Reinhardt, 2012b, S. 75), somit „die Bewegung der Sache Politik (bzw. Gesellschaft, Wirtschaft, Recht) und die Bewegung des Lernens [vermitteln] (sie bauen Brücken, über die Lernenden gehen können)“ (ebd.). Sander (2008) vergleicht didaktische Prinzipien mit „Scheinwerfern, die Lernmöglichkeiten in der Auseinandersetzung mit fachlich bedeutsamen Problemfeldern jeweils unterschiedlich ausleuchten“ (S. 191) und bezeichnet sie als „begriffliche Tools für die Auswahl und didaktische Strukturierung von Lerngegenständen“ (ebd., S. 189). Es gibt unterschiedliche Auffassungen dazu, was ein als solches zunächst fachunspezifisches didaktisches Prinzip (Detjen, 2007, S. 319) zu einem politikdidaktischen Prinzip macht. So gilt nach Reinhardt (2012b), „[e]in Prinzip ... nur dann als fachdidaktisches Prinzip, wenn sich ihm mindestens eine Methode zuordnen lässt“ (S. 75), wobei diese Schlussfolgerung von Detjen (2007, S. 321) am Beispiel des exemplarischen Lernens infrage gestellt wird. Alternative Systematisierungen sollen Prinzipien danach unterscheiden, ob sie „eher bei der *Sache* oder eher bei den *Lernenden* angesiedelt sind“ (ebd.) bzw. „ob sie eher auf die *Auswahl* und *Strukturierung* der Lerngegenstände oder eher auf die Art und Weise der *Vermittlung* der Lerngegenstände gerichtet sind“ (ebd.). Auch Sander (2008, S. 190) betont die Möglichkeit der analytischen Trennung von didaktischen Prinzipien, macht gleichzeitig aber auch auf deren Überschneidungs- und Ergänzungsmöglichkeiten aufmerksam; Detjen (2007, S. 321) sieht diesen Kompatibilitätsansatz nicht vollumfänglich gegeben und weist auf mögliche Unvereinbarkeiten zwischen den Prinzipien Schülerorientierung und Wissenschaftsorientierung hin. Ungeachtet der unterschiedlichen Klassifikationsansätze wird betont, dass sich der ‚Wert‘ von didaktischen Prinzipien und zugeordneten Methoden an ihrem oben beschriebenen Werkzeugcharakter messen lassen muss, also daran, wie gut sie bei der Zielerreichung des Unterrichts helfen können (ebd.; Manzel & Neumann, 2015, S. 77), was Manzel (2014) so zusammenfasst:

„Eine fachdidaktische Ausbildung erschöpft sich nicht in der Darbietung zahlreicher fachdidaktischer Prinzipien und Methoden für den Unterricht, sondern zeigt theoretische Begründungszusammenhänge, warum welches Prinzip und welche Methode für welches Ziel und welchen Lehr-Lernprozess besonders geeignet erscheint, und weist diese in empirischen Studien nach.“ (S. 62).

Da empirische Forschungen zur Wirksamkeit von fachdidaktischen Prinzipien zurzeit in der Politikdidaktik weiterhin ausstehen, wird „[i]hr Status ... [als] der von *Klugheitsprinzipien*“ (Detjen, 2007, S. 321) bezeichnet. Fachdidaktische Prinzipien können daher Lehrer/-innen bei der Auswahl und Begründung von Methoden auf Mikro-, Makro- und Mesoebene des

Unterrichts (Meyer, 2011, S. 74ff.) unterstützen, jedoch nicht die Ziele und Inhalte des Unterrichts konstituieren (Manzel & Neumann, 2015, S. 77).

Zur Übersicht über die Konzeptualisierung der Facetten des fachdidaktischen Wissens in unterschiedlichen Studien lässt sich zusammenfassend festhalten, dass unter dem Wissen über Instruktions- und Vermittlungsstrategien mehrheitlich adäquate „Repräsentations- und Erklärungsformen“ (Weschenfelder, 2014a, S. 169) und unter dem Wissen über Schülerkognitionen unter anderem typische (Fehl-)Vorstellungen der Lernenden verstanden werden. Die vorliegenden domänenspezifischen Ausformulierungen hierzu sind dabei unterschiedlich konkret. So wird in den Naturwissenschaften beispielsweise der Umgang mit Experimenten betont, was sich auch aus empirischen Forschungen zu deren Einsatz im Unterricht ableiten lässt. Da in der Politikdidaktik Untersuchungen zur Wirksamkeit von bestimmten fachspezifischen Methoden und Prinzipien derzeit noch ausstehen, werden die besonderen Herausforderungen bei der Konzeptualisierung und Operationalisierung des fachdidaktischen Wissens betont (ebd.). Hierbei wird die Notwendigkeit angeführt, auch auf ausgewählte normative politikdidaktische Referenzpunkte zu rekurrieren; ein entsprechender Diskurs lässt sich in der Politikdidaktik unter anderem zu den Funktionen und Wirkungsweisen von fachdidaktischen Prinzipien ausmachen, wobei der ‚Werkzeugcharakter‘ dieser Prinzipien im Hinblick auf den Outcome des Schulunterrichts betont wird (Manzel & Neumann, 2015, S. 77).

Die Konzeptualisierung der Facette des Wissens über Schülerkognitionen unterscheidet sich zwischen den Fächerdomänen sowie innerhalb dieser weniger deutlich. Im Wesentlichen teilen die referierten Untersuchungen die Auffassung, dass Lehrkräfte über Wissen zu typischen (Fehl-)Vorstellungen von Lernenden verfügen müssen, um den Unterricht im Planungsprozess sowie in der eigentlichen Durchführung an deren Kognitionen anpassen zu können. Die Domänenspezifität kommt in dieser Kompetenzfacette noch deutlicher zum Ausdruck als im Wissen über Instruktions- und Vermittlungsstrategien, weshalb die unterschiedlichen Ansätze insbesondere aus einer fächerübergreifenden Perspektive lediglich auf dem in Tabelle 2 aufgeführten basalen Konzeptualisierungsniveau miteinander verglichen werden können. Hierzu bleibt aus domänenspezifischer Sicht festzuhalten, dass es in der Politikdidaktik weiterhin Forschungsdesiderata zu typischen Schwierigkeiten und Fehlern von Schülerinnen und Schülern gibt. Im Hinblick auf das Verhältnis der Facette des

fachdidaktischen Wissens zu den diagnostischen Fähigkeiten von Lehrkräften wird für die nachfolgenden Überlegungen in dieser Arbeit in Anlehnung an die Konzeptualisierung in COACTIV (Brunner et al., 2011) davon ausgegangen, dass sich diagnostische Fähigkeiten als Aspekt des fachdidaktischen Wissens insbesondere im Wissen über (Fehl-)Vorstellungen der Lernenden zeigen (vgl. auch D. Neumann, 2014); somit bleibt der pädagogisch-psychologische Aspekt der Leistungsbeurteilung aufgrund des Fokus der vorliegenden Arbeit ausgespart (Brunner et al., 2011, S. 217).

2.2.2 Strukturdimension *Inhaltsbereiche*

Schon das einleitend in Kapitel 2.2 angesprochene Verhältnis von Fachwissen und fachdidaktischem Wissen verdeutlicht, dass sich dieses Wissen im Sinne der Taxonomie Shulmans (1986, 1987) „auf die *inhaltsbezogenen* (nicht allgemein-methodischen) unterrichtlichen Aufgaben“ (Krauss et al., 2008b, S. 227; vgl. auch Weschenfelder, 2014a, S. 109) bezieht und insofern die Frage nach der ‚inhaltlichen Füllung‘ der Facetten des fachdidaktischen Wissens nicht vollständig beantwortet werden kann, ohne zu spezifizieren, auf welche Inhalte sich dieses Wissen bezieht. In einigen Modellierungen des Professionswissens sowie in Teilmodellierungen des fachdidaktischen Wissens von Lehrkräften wird das Wissen daher nicht nur domänenspezifisch, sondern auch inhaltspezifisch (van Driel, Verloop & de Vos, 1998; Schmelzing, 2010, S. 31; Schmelzing et al., 2010, S. 191) bzw. themenspezifisch (Blömeke et al., 2008; Tepner et al., 2012, S. 9; van Dijk & Kattmann, 2010, S. 8) konzeptualisiert. So treten im mehrdimensionalen Modell des Professionswissens im Projekt *ProwiN* (*Professionswissen von Lehrkräften in den Naturwissenschaften*) (Tepner et al., 2012) neben die kognitiven Aktivitäten (Wissensarten; vgl. auch Kapitel 2.2.3) sowie die Darstellung der (operationalisierbaren) Facetten des fachdidaktischen und des pädagogischen Wissens Inhaltsbereiche, die wegen der fächerübergreifenden Modellierung nicht näher spezifiziert sind. Dies wird damit begründet, dass „[g]rundsätzlich ... von einer Domänenspezifität des Fachwissens auch innerhalb eines Faches auszugehen [ist], sodass hohes Wissen in einem Themengebiet nicht automatisch hohes Wissen in einem anderen Themengebiet impliziert“ (Tepner et al., 2012, S. 9) (Abbildung 6).

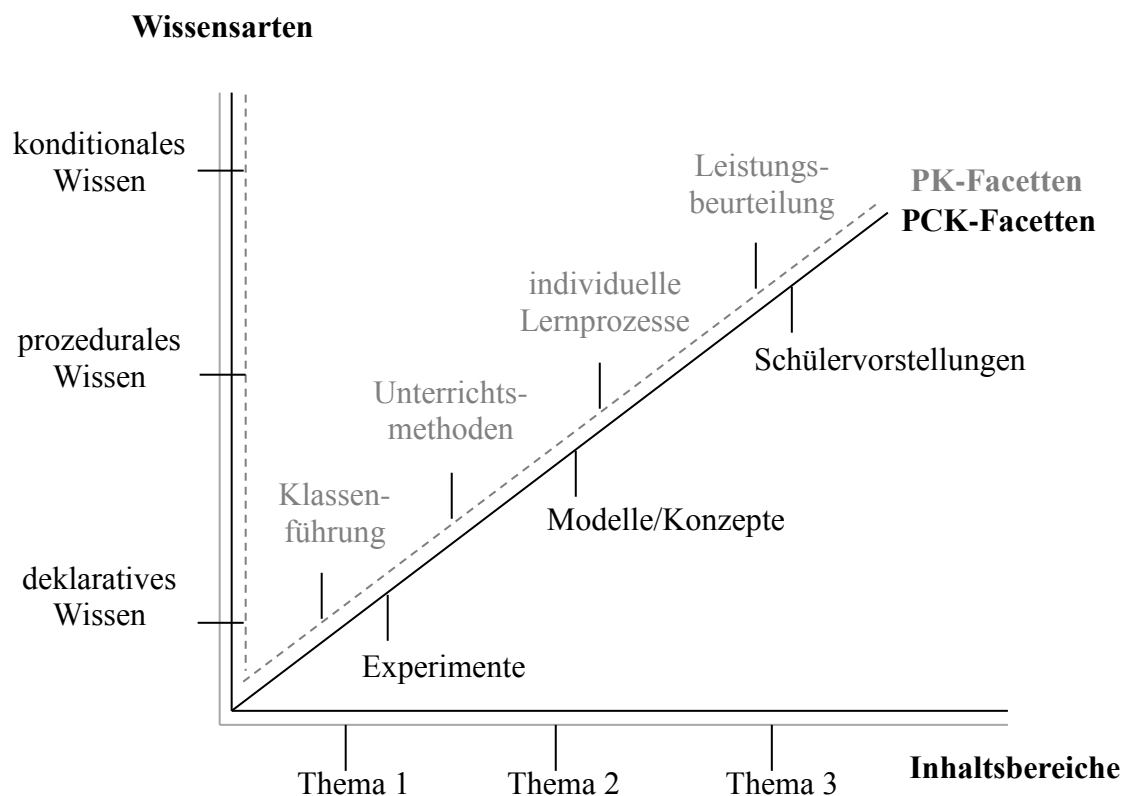


Abbildung 6: Modell des Professionswissens von Lehrkräften in den Naturwissenschaften im Projekt ProwiN (eigene Darstellung in Anlehnung an Tepner et al., 2012, S. 19)

Einen ähnlichen Konzeptualisierungsansatz verfolgen Blömeke et al. (2008) für das mathematisch-fachdidaktische Wissen in der MT21-Studie in den drei Dimensionen *kognitive Aktivität*, *Anforderungen* und *Inhaltsbereiche*, wobei letztere aufgrund der Fachspezifität des Modells unmittelbar benannt sind. Auch in Modellierungen, die sich exklusiv auf das in diesem Kapitel interessierende fachdidaktische Wissen konzentrieren, erfolgt die Konzeptualisierung teilweise entlang von Inhaltsbereichen, was Gramzow et al. (2013) im Rahmen ihrer Modellierung des fachdidaktischen Wissens von angehenden Physiklehrkräften nach einem Vergleich unterschiedlicher Konzeptualisierungsansätze so zusammenfassen:

„Wie oben beschrieben, bezieht sich die Inhalts-Dimension, sofern vorhanden, auf konkrete Inhalte des jeweiligen Faches und ist damit über verschiedene Konzeptualisierungen hinweg gut vergleichbar. Die Notwendigkeit einer solchen Dimension liegt darin begründet, dass es keine hinreichend belastbaren Indizien für inhaltsunabhän-

gige fachdidaktische Wissensbereiche (die ohne weiteres von einem physikdidaktischen Inhalt auf einen anderen übertragen werden können) gibt.“ (S. 13).

Bei Schmelzing (2010) sowie Schmelzing et al. (2010) enthält die mehrdimensionale Modellierung des biologiedidaktischen Wissens ebenfalls eine inhaltsbezogene Strukturdimension; diese umfasst „die fachlichen *Inhalte* (Shulman, 1987; Shulman, 1986; Veal & Kubasko, 2003), auf die sich die fachdidaktischen Kenntnisse und Fertigkeiten von Lehrkräften beziehen“ (Schmelzing, 2010, S. 29) (Abbildung 7).

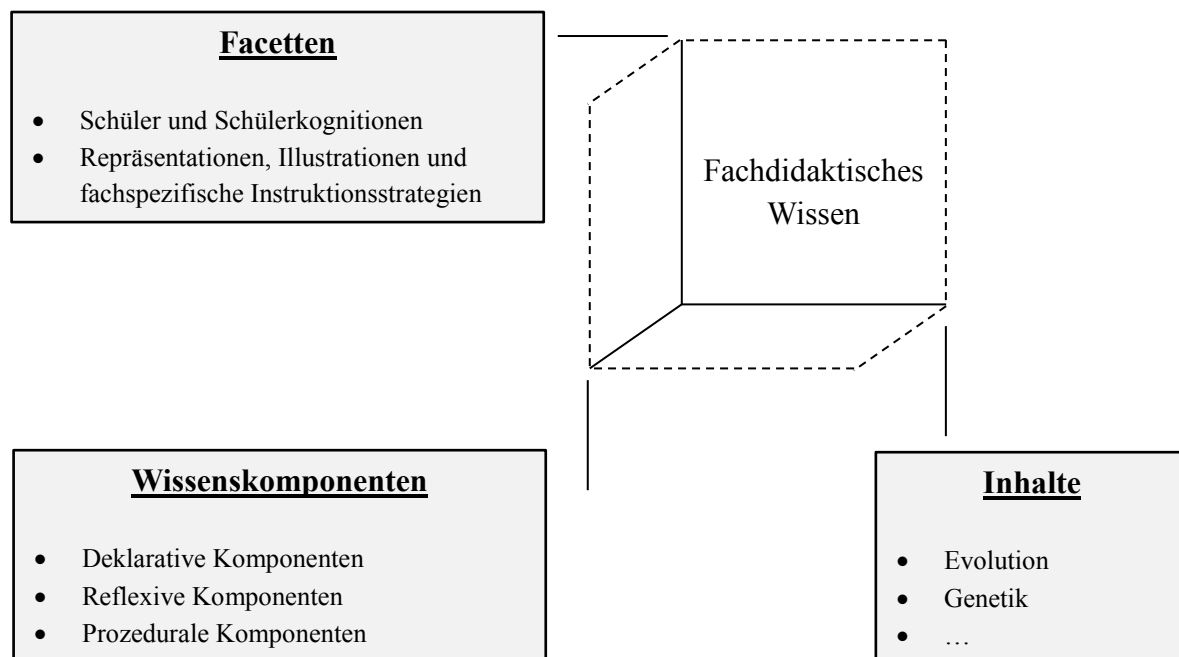


Abbildung 7: Heuristisches Strukturmodell zum biologiedidaktischen Wissen (eigene Darstellung nach Schmelzing, 2010, S. 29)

Die in diesem heuristischen Modell dargestellten beispielhaften Inhalte der Biologie, auf die sich das fachdidaktische Wissen von Lehrkräften beziehen soll, werden schließlich bei Schmelzing im Zuge der Testskalenentwicklung anhand des exemplarischen Inhaltsbereiches „Blut und Blutkreislaufsystem des Menschen“ (ebd., S. 45) operationalisiert. Auch in einem weiteren Ansatz zum biologiedidaktischen Wissen von Lehrkräften – dem *ERTE*-Modell (*Educational Reconstruction for Teacher Education*) von van Dijk und Kattmann (2007) – erfolgt eine Konzeptualisierung entlang einer inhalts- bzw. themenspezifischen Ausdifferenzierung, da kritisiert wird, dass im Laufe der Weiterentwicklung der Kategorie des fachdi-

daktischen Wissens „das wichtigste Merkmal des PCK-Konzepts, der Themenbezug, verloren gegangen [ist] (vgl. Hashweh, 2005)“ (van Dijk & Kattmann, 2010, S. 8). Die Forschungsfragen, die van Dijk und Kattmann für die empirische Überprüfung der Modellannahmen zum biologiedidaktischen Wissen mit dem exemplarischen Fokus auf das Thema *Evolution* stellen, werden infolgedessen so formuliert:

- „(1) Über welche ‚fachlichen Kenntnisse für das Lehren‘ von Evolution verfügen die Lehrenden?
- (2) Welche Kenntnisse von Schülervorstellungen haben Lehrende zum Thema Evolution?
- (3) Welche Kenntnisse von den verschiedenen Instruktionsformen haben Lehrende zum Thema Evolution?“ (Ebd., S. 9)

Der Ansatz einer inhaltsspezifischen Ausdifferenzierung lässt sich auch bei der Konzeptualisierung des politisch-fachdidaktischen Wissens in der PKP-Studie (Ober et al., 2012) erkennen. Es wird dort unter Rekurs auf Baumert und Kunter (2006) davon ausgegangen, dass das Fachwissen die Grundlage ist, „*auf der fachdidaktische Beweglichkeit entstehen kann*“ (S. 496; vgl. auch Weschenfelder, 2014a, S. 107). Gleichzeitig wird festgestellt, dass es „[a]ufgrund der diffusen Systematik des Schulfaches Politik ... jedoch schwierig [ist,] Inhalte zu identifizieren (vgl. Eck & Weißen, 2009, S. 25)“ (Weschenfelder, 2014a, S. 107). In der PKP-Studie wird daher Politik als Kern (Massing & Weißen, 1995) des Modells angesehen und zu diesem Zweck die Empfehlung eines Kerncurriculums für politikwissenschaftliche Studiengänge der Deutschen Vereinigung für Politische Wissenschaft (DVPW, 2003) herangezogen (Weschenfelder, 2014a, S. 163f.). Diese „beinhalten die fünf Bereiche Politische Theorie ..., Innenpolitik und politisches System der Bundesrepublik Deutschland (BRD), Politische Systeme im Vergleich (VR), Außenpolitik und Internationale Beziehungen (Außenpol./IB) sowie Methoden der Politikwissenschaft (Methoden)“ (ebd., S. 164) und sind in Verbindung mit – aus einer Analyse von politikdidaktischen Lehramtsstudienfächern deduktiv abgeleiteten – politikwissenschaftlichen und politikdidaktischen Studienbereichen Grundlage für die Operationalisierung der Testkonstrukte (ebd., S. 165ff.). Die Itementwicklung in der Dimension des Fachwissens erfolgt mit den Schwerpunkten „politisches System der Bundesrepublik Deutschland, Politische Systeme im Vergleich und Internationale Beziehungen“ (ebd., S. 167); es wird davon ausgegangen, dass sich die Inhaltsspezifität des politikdidaktischen Wissen vor allem im „Wissen um den Aufbau von

Konzepten bei Schüler/-innen, um alternative Repräsentations- und Erklärungsformen und das Erkennen von Fehlkonzepten [zeigt]“ (ebd., S. 169).

Bis hierher lässt sich zusammenfassen, dass einige Modelle unterschiedliche domänenspezifische Inhalts- bzw. Themenbereiche ausdifferenzieren, auf die sich das konzeptualisierte Wissen bezieht. Im Rahmen der PKP-Studie (Oberle et al., 2012) findet ebenfalls eine inhaltspezifische Ausdifferenzierung statt (Weschenfelder, 2014a, S. 167f.). Eine solche Spezifizierung kann theoretisch begründet (Shulman, 1986, 1987; van Dijk & Kattmann, 2010, S. 8; vgl. auch Kapitel 2.2) und durch empirische Befunde gestützt werden (vgl. u. a. Tepner et al., 2012, S. 9). Wenn im Folgenden von einer Konzeptualisierung des politikdidaktischen Wissens entlang einer Strukturdimension *Inhalte* ausgegangen wird, erfolgt dies in Anlehnung an van Dijk und Kattmann (2010, S. 9) in der Annahme, dass Lehrkräfte neben fachbezogenen Kenntnissen entsprechendes Wissen darüber haben müssen, welche *inhaltsbezogenen* (Fehl-)Vorstellungen der Lernenden zu erwarten sind; ebenso wird von einem *inhaltsbezogenen* Wissen über Instruktions- und Vermittlungsstrategien ausgegangen. Mit dieser Strukturdimension soll ein weiterer theoretischer Referenzpunkt für die im späteren Verlauf der Arbeit zu leistende Entwicklung der Testinstrumente zur Erhebung des politikdidaktischen Wissens (Kapitel 6.3) begründet werden. Gleichzeitig soll eine für den explorativen Zweck der vorliegenden Arbeit notwendige Fokussierung ermöglicht werden.

Die Rekonstruktion des fachdidaktischen Zugangs im Rahmen der Strukturdimension *Inhaltsbereich* erfolgt in dieser Arbeit zum Thema *Parteienverbote in einer Parteiendemokratie* (vgl. auch Kapitel 6.3) und deutet damit auf eine didaktische Zugangsweise nach dem (fach-)didaktischen *Kontroversitätsprinzip* hin. Diese eher an Kriterien der unterrichtlichen Anforderungssituation orientierte Begründung zur inhaltsbezogenen ‚Füllung‘ des politikdidaktischen Wissens kann durch eine normativ-deduktive Sichtweise ergänzt werden, da das Kontroversitätsprinzip als „Generalprinzip des Politikunterrichts“ (Reinhardt, 2012b, S. 75) angesehen wird, welches „zum Kern der Berufsethik des politischen Bildners [zählt]“ (Grammes, 2005, S. 126). In der Politikdidaktik wird dies durch den Umstand begründet, dass politische Bildung sich grundsätzlich in einem Spannungsfeld zwischen zwei ideologischen Ausrichtungen befindet: der Erziehung zum Staat und der Bildung des Menschen (ebd., S. 127). Vor 1945, zur Zeit des Kaiserreiches und des Nationalsozialismus, hatte politische Bildung vor allem einen instrumentell-herrschaftslegitimierenden Charakter und damit

die Aufgabe, dass „[e]in bestehender gesellschaftlich-politischer Zustand ... im Interesse der von ihm profitierenden Machtgruppen durch politische Erziehung legitimiert und vor Kritik geschützt werden [soll]“ (Sander, 2005, S. 15). Nach dem Zweiten Weltkrieg setzte sich im Laufe der demokratischen Neuorientierung in Deutschland der Anspruch an politische Bildung durch, sie solle zur Mündigkeit der Schüler/-innen beitragen. Nach diesem Ansatz „dient politische Bildung einer eigenständigen Auseinandersetzung der Lernenden mit dem Wirklichkeitsbereich Politik, ohne die Ergebnisse dieser Auseinandersetzung, die politischen Meinungen, Urteile und Überzeugungen, zu denen die Lernenden im Einzelnen kommen können, vorwegnehmen zu wollen“ (ebd., S. 17). Die im Beutelsbacher Konsens 1976 formulierten drei Grundsätze reflektieren dieses Selbstverständnis politischer Bildung und werden von Wehling (1977) so zusammengefasst:

„1. *Überwältigungsverbot*. Es ist nicht erlaubt, den Schüler – mit welchen Mitteln auch immer – im Sinne erwünschter Meinungen zu überrumpeln und damit an der ‚Gewinnung eines selbstständigen Urteils‘ zu hindern Hier genau verläuft die Grenze zwischen Politischer Bildung und *Indoktrination*. Indoktrination aber ist unvereinbar mit der Rolle des Lehrers in einer demokratischen Gesellschaft und der – rundum akzeptierten – Zielvorstellung von der Mündigkeit des Schülers.

2. Was in Wissenschaft und Politik *kontrovers* ist, muß auch im Unterricht kontrovers erscheinen. Diese Forderung ist mit der vorgenannten aufs engste verknüpft, denn wenn unterschiedliche Standpunkte unter den Tisch fallen, Optionen unterschlagen werden, Alternativen unerhört bleiben, ist der Weg zur Indoktrination beschritten.

3. Der Schüler muß in die Lage versetzt werden, eine *politische Situation* und seine *eigene Interessenlage zu analysieren*, sowie nach Mitteln und Wegen zu suchen, die vorgefundene politische Lage im Sinne seiner Interessen *zu beeinflussen*.“ (S. 179f.)

Der in der Politikdidaktik allgemein akzeptierte Beutelsbacher Konsens (Sander, 2005, S. 18) soll bis heute Orientierungshilfe und gleichzeitig ‚Mahner‘ in dem bereits skizzierten Konflikt sein, in dem sich schulische politische Bildung grundsätzlich befindet: Sie orientiert sich „an einem Verständnis des Menschen als Subjekt, dessen Mündigkeit im Sinne selbstständigen Urteilens und Handelns ... [sie] fördern will“ (Sander, 2008, S. 35), muss gleichzeitig aber als Unterrichtsfach eine Herrschaftslegitimationsfunktion der Institution Schule miterfüllen (Fend, 2008, S. 46f.). Die *Rahmenvorgabe Politische Bildung* (MSWF, 2001) für Nordrhein-Westfalen stellt im Zusammenhang mit der Darstellung zentraler Leitideen schulischer politischer Bildung Folgendes fest:

„Bei Wertvorstellungen, die in der politischen Diskussion zumeist als ‚Grundwerte‘ bezeichnet werden – wie Freiheit, Frieden, Gerechtigkeit – ist im Unterricht immer zu bedenken, dass ihre inhaltliche Konkretisierung, ihre Gewichtung und ihr Verständnis zueinander selbst schon kontroverse Gegenstände der politischen Diskussion sind. Die Politische Bildung darf dies nicht stellvertretend für die Schülerinnen und Schüler vorentscheiden. Dennoch wird die Politische Bildung auch durch zentrale Wertvorstellungen bestimmt, die als normative Grundorientierungen demokratischer politischer Bildung die Grundlage für die konkreten, im Unterricht zu erreichenden Ziele bilden.“ (S. 14)

Ein sich in der freiheitlich-demokratischen Grundordnung widerspiegelnder Wertekonsens zählt also „zu den nicht-kontroversen Rahmenbedingungen für eine demokratische Gesellschaft“ (Grammes, 2005, S. 129), wobei die Aushandlung dessen, was kontrovers ist, selbst wieder Gegenstand von Auseinandersetzungen ist (ebd.).

Die unterrichtliche Umsetzung des Kontroversitätsprinzips stellt hohe Anforderungen an das fachdidaktische Wissen und Können von Politiklehrkräften. Sie müssen über diagnostische Fähigkeiten (vgl. auch Kapitel 2.2.1) verfügen, um feststellen zu können, ob es sich bei den Schülerinnen und Schülern um eine „*argumentationshomogene*“ (Detjen, 2007, S. 328), eine „*argumentationsheterogene*“ (ebd.) oder eine „*apathisch-indifferente* Lerngruppe“ (ebd.) handelt, und dann entscheiden, welche Rolle sie beispielsweise in Diskussionen einnehmen sollten: die des Provokateurs, des Moderators oder des Animateurs (ebd.). Weiterhin muss die Lehrkraft fachdidaktisches Wissen über den Einsatz von Unterrichtsmethoden und -medien haben, die sich dazu eignen, das Prinzip unterrichtlich wirksam werden zu lassen, wobei „Nachbildungen öffentlicher Gesprächsformen ... dem Kontroversitätsprinzip methodisch besonders gerecht [werden]“ (ebd., 329). In diese Kategorie fallen unter anderem „die Pro-Contra-Debatte, Podiumsdiskussion, Fishbowl, Talkshow, Tribunal“ (Grammes, 2005, S. 134). Hinsichtlich des Einsatzes von Medien wird Kontroversität in einem doppelten Sinne bedeutsam. So muss die Auswahl der Medien dem Kontroversitätsgebot gerecht werden, indem beispielsweise unterschiedliche Quellen verwendet werden, um den Lernenden einen multiperspektivischen Zugang zu ermöglichen (ebd., S. 134f.). Gleichzeitig sollen im Unterricht Medien authentisch dargeboten werden, damit unterschiedliche Perspektiven unverfälscht miteinander verglichen werden können (ebd., S. 135).

Wenn der Blick auf die relevanten Rahmenvorgaben und Schulcurricula gerichtet wird, zeigt sich auch dort die Rezeption von Kontroversität sowohl im Bereich der Lernzielerwartungen als auch der inhaltlichen Ausgestaltung des Unterrichts. So betont die *Rahmenvorgabe Poli-*

tische Bildung (MSWF, 2001) für Nordrhein-Westfalen die grundlegende Bedeutung der Kontroversität als Bestandteil des Beutelsbacher Konsenses unter anderem für die Wahl und Ausgestaltung der Themen, die „offene Fragen, unentschiedene Konflikte und kontrovers beurteilte Sachverhalte aus Politik und Gesellschaft entsprechend aufgreifen [sollen]“ (S. 23). Eine Durchsicht der Kernlehrpläne für die Unterrichtsfächer *Gesellschaftslehre* (Hauptschule und Gesamtschule), *Politik* (Realschule) sowie *Politik/Wirtschaft* (Gymnasium) für die Sekundarstufe I zeigt, dass Kontroversität in den Lernzielerwartungen als Facetten der Dimensionen Urteilskompetenz und Handlungskompetenz durchgehend aufgeführt ist (MSW, 2007, 2011, o. J. a, o. J. b). Schüler/-innen des Gymnasiums sollen „zu einer selbstständigen, begründeten und kriterien- oder kategorienorientierten Beurteilung politischer Ereignisse, Probleme und Kontroversen“ (MSW, 2007, S. 19) befähigt werden. Diese Lernziele finden sich in den Kernlehrplänen der Sekundarstufe I durchgängig im Bereich der Urteils- sowie der Handlungskompetenz (MSW, 2007, 2011, o. J. a, o. J. b). Ähnliche Zielvorstellungen enthält auch der Kernlehrplan für die Sekundarstufe II (Gymnasium und entsprechende Jahrgangsstufen der Gesamtschule) für die Fächer *Sozialwissenschaften* bzw. *Sozialwissenschaften/Wirtschaft*, wobei vor allem die zu bearbeitenden Inhaltsfelder noch umfassender auf die Wichtigkeit einer kontroversen unterrichtlichen Aufbereitung politischer und gesellschaftlicher Problemlagen hinweisen (MSW, 2014a).

2.2.3 Strukturdimension Wissensarten

Die Auffassung der Literatur, dass Lehrer/-innen neben dem Wissen über unterrichtsrelevante Fakten und Konzepte (Wissen, dass ...) auch Kenntnisse darüber haben müssen, in welcher Art und Weise dieses Wissen in unterrichtlichen Anforderungssituationen in eine angemessene Handlungsstrategie überführt werden kann (Wissen, wie ...; vgl. auch Kapitel 2.1.2), spiegelt sich in den Konzeptualisierungen des fachdidaktischen Wissens wider. Eine Unterscheidung in deklaratives und prozedurales Wissen lässt sich in vielen Modellen zum fachdidaktischen Wissen identifizieren, wobei es zum Teil unterschiedliche Auffassungen darüber gibt, was speziell das prozedurale fachdidaktische Wissen kennzeichnet. So unterscheiden Schmelzing et al. (2010) für das biologiedidaktische Wissen folgende Wissensarten:

„Deklaratives fachdidaktisches Wissen beschreibt explizites, d.h. verbalisierbares, Begriffswissen (*Wissen, dass...*), deklarativ-kausales Begründungswissen (*Wissen, warum...*), sowie

deklarativ-konditionales Bedingungswissen (*Wissen, wann und wozu...*; Kaufhold, 2006; Polanyi, 1958; Ryle, 1949; Schelten, 2004; vgl. auch Paris, Lipson & Wixson, 1983). Das prozedurale fachdidaktische Wissen umfasst Können (*Wissen, wie...*), welches in Form von überwiegend impliziten, d.h. nicht verbalisierbaren, und automatisierten Handlungsrouninen vorliegt (Anderson, 1981; Polanyi, 1958; Ryle, 1949).“ (S. 192)

Diese Ausdifferenzierung in ein eher auf die Aktivierung von Fakten und Begriffen ausgerichtetes deklaratives fachdidaktisches Wissen und ein prozedurales fachdidaktisches Wissen, das „mehrheitlich implizites und handlungsbezogenes Verfahrenswissen, automatisierte Handlungsrouninen sowie peri-aktionale, das heißt simultan zur Handlung vorliegende, Wahrnehmungs- und Verarbeitungsroutinen [umfasst] (Baumert et al., 2004; Baxter & Lerman, 1999; Park & Oliver, 2008; Tamir, 1988; van Driel et al., 1998; Weinert et al., 1990)“ (Schmelzing, 2010, 31), lässt die Betonung einer angenommenen „Handlungswirksamkeit des Professionswissens“ (Gramzow et al., 2013, S. 17) erkennen. Reflexives Wissen bezeichnet bei Schmelzing (2010) „prä-aktionale, das heißt vor der Handlung vorliegende, Planungsfertigkeiten sowie peri- und post-aktionale, das heißt nach der Handlung vorliegende Beobachtungs-, Bewertungs-, Analyse [*sic*] und Reflexionsfertigkeiten (Lindmeier & Heinze, 2008; Park & Oliver, 2008; Schön, 1983; Rambow & Bromme, 2000)“ (S. 31). In einem ähnlichen Verständnis ist die *reflexive Kompetenz* bei Lindmeier et al. (2012) konzeptualisiert. Um „z. B. schnelle, spontane Reaktion vs. vertiefende Reflexion“ (ebd., S. 104) zu unterscheiden, werden „[u]nter reflexiver Kompetenz ... hier domänenspezifische professionelle Fähigkeiten zusammengefasst, die bei der Vor- und Nachbereitung von Unterricht benötigt werden“ (ebd.). Tepner et al. (2012) merken im Hinblick auf diese Wissenskomponente an, dass die „vorgeschlagene Unterteilung des fachdidaktischen Wissens in deklaratives, prozedurales und reflexives Wissen ... primär der Verbindung von Wissen und Handeln Rechnung [trägt], indem mittels Videovignetten das reflexive Wissen in Bezug auf bestimmte Unterrichtssituationen gemessen wird (Schmelzing, Wüsten, Sandmann & Neuhaus, 2010)“ (S. 17) und „[d]as reflexive Wissen ... somit retrospektiv nach dem Handeln abgerufen [wird]“ (ebd.).

Eine Unterscheidung in deklarative und prozedurale Wissenskomponenten, obgleich ohne weitere Ausdifferenzierung des deklarativen Wissens wie bei Schmelzing (2010) bzw. Schmelzing et al. (2010), findet sich bei Riese und Reinhold (2010) für das fachdidaktische Wissen von angehenden Physiklehrkräften. Deklaratives Wissen als „Kenntnis von Konzepten und Fakten“ (ebd., S. 172) wird dort unterschieden vom prozeduralen Wissen, „welches

sich beispielsweise in der Verfügbarkeit von Handlungsrouinen äußert“ (ebd.). Beide Ansätze teilen die Auffassung, dass sich prozedurales Wissen daher nicht nur auf Handlungsdispositionen bezieht, sondern auch auf Handlungsrouinen (Gramzow et al., S. 17). Bei Tepner et al. wird das deklarative und prozedurale Wissen in Anlehnung an die Konzeptualisierung von Paris, Lipson und Wixson (1983) um das konditionale Wissen ergänzt, welches beschrieben wird als „das *Wissen*, wann und warum ein Prozess oder eine Handlung angewandt werden kann oder sollte“ (Tepner et al., 2012, S. 17; vgl. auch Abbildung 6). Demgegenüber ist „[u]nter dem prozeduralen Wissen ... die Kenntnis von (Unterrichts-)Handlungen und -prozessen als dem *Wissen*, wie etwas abläuft, zu fassen“ (ebd.). Im Modell von Tepner et al. (ebd., S. 19) lassen sich die Facetten des fachdidaktischen Wissens allen drei Wissensarten zuordnen.

Im Unterschied dazu wird in der eindimensionalen Modellierung von Riese (2009) der Zusammenhang von Wissensarten und Facetten fachdidaktischen Wissens „auf einem Kontinuum von ‚eher deklarativ‘ zu ‚eher prozedural‘ verortet“ (Gramzow et al., 2013, S. 12) (Abbildung 8).

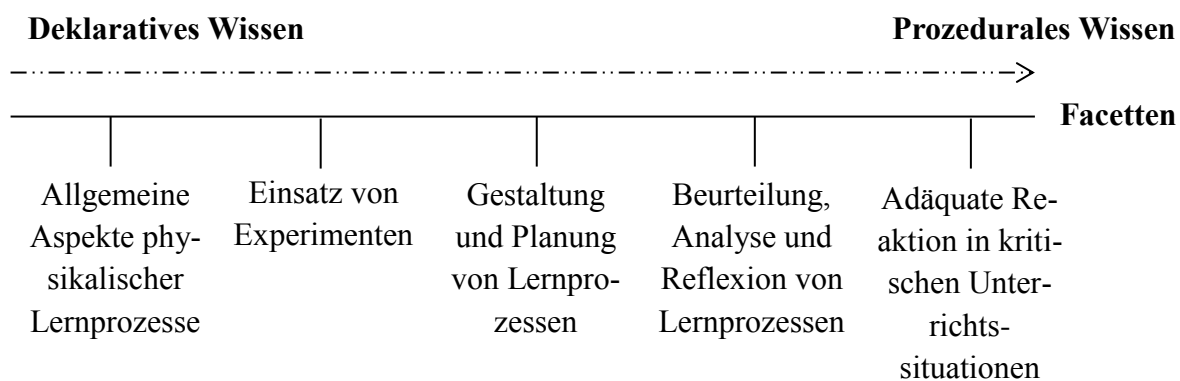


Abbildung 8: Modell physikalisch-fachdidaktischen Wissens bei Riese (2009) (eigene Darstellung nach Gramzow et al., 2013, S. 13)

Diese Modellierung stellt den Anwendungsbezug des prozeduralen Wissens im Sinne eines konkreten unterrichtlichen Handelns stärker in den Vordergrund als das Modell von Tepner et al. (2012, S. 17), das mehr den Charakter des prozeduralen Wissens als kognitive Handlungsdisposition betont.

Es lässt sich resümieren, dass das fachdidaktische Wissen in den referierten Ansätzen mindestens in die beiden Arten des deklarativen und des prozeduralen Wissens unterteilt wird. Während unter dem deklarativen fachdidaktischen Wissen mehrheitlich ein Wissen über didaktische Sachverhalte und Fakten gefasst wird, variiert die Vorstellung zum prozeduralen fachdidaktischen Wissen zwischen einem Wissen *über* mögliche Handlungsweisen, einem Wissen im Sinne eines Verfügenkönnens über Handlungen und einem Wissen, das auch Handlungsrouinen umfasst. Insgesamt „spiegelt [dies] die Absicht der Modelle wieder [*sic*], Professionswissen als Teil von Handlungskompetenz zu modellieren und dabei Bezüge zum Können oder direkt zum Handeln herzustellen“ (Gramzow et al., 2013, S. 18). Auf die Verortung des fachdidaktischen Wissens im Rahmen übergeordneter Modelle professioneller Handlungskompetenzen soll daher im nachfolgenden Kapitel eingegangen werden.

2.3 Politikdidaktisches Wissen und Können als Aspekt professioneller Handlungskompetenzen

Das professionelle deklarative und prozedurale Wissen wird in kompetenztheoretischen Ansätzen in der Regel als Aspekt professioneller Handlungskompetenzen von Lehrpersonen konzeptualisiert und durch weitere moderierende Variablen wie professionelle Überzeugungen und motivationale Aspekte ergänzt (Weschenfelder, 2014a, S. 11). In diesem Kapitel wird zunächst der Kompetenzbegriff aus einer kognitionspsychologischen Perspektive, wie er in der vorliegenden Arbeit rezipiert wird, erläutert (Kapitel 2.3.1). Anschließend wird auf die Grundlagen der Modellierung professioneller Handlungskompetenzen eingegangen (Kapitel 2.3.2) und es werden Ansätze einer domänenspezifischen Kompetenzdiagnostik aufgezeigt (Kapitel 2.3.3), bevor schließlich in Kapitel 2.3.4 Überlegungen zum professionellen Wissen und Können von Politiklehrkräften verschiedener Verfasser/-innen im fachdidaktischen Diskurs dargestellt und miteinander verglichen werden.

2.3.1 Zum Kompetenzbegriff aus einer kognitionspsychologischen Perspektive

Wie das Angebots-Nutzungs-Modell unterrichtlicher Wirkungen (Helmke, 2014, S. 71) verdeutlicht, reichen allgemeine kognitive Fähigkeiten von Lehrpersonen nicht aus, um Lehr-Lern-Situationen adäquat planen und gestalten zu können, weshalb der Kompetenzbegriff

auch „als Gegenbegriff zu generalisierten, kontextunabhängigen kognitiven Leistungskonstrukten eingeführt [wurde], wie sie für die Intelligenzforschung und -diagnostik typisch sind“ (Klieme et al., 2007, S. 6). Eine breit rezipierte Auffassung von Kompetenzen, nach der diese als komplexe Konstrukte definiert werden, die Wissen und Fähigkeiten, motivationale Orientierungen sowie Fähigkeiten der Selbstregulation umfassen, hat dabei Weinert (2001, S. 27f.) geprägt. In einem solchen kognitionspsychologischen Verständnis sind Kompetenzen „in der Regel kontextspezifisch und handlungsbezogen definiert“ (Frey & Hartig, 2013, S. 54) sowie prinzipiell erlernbar und intern graduierbar (Klieme et al., 2007/2009, S. 65). Es gibt weitere Definitionen, in denen wissenschaftstheoretisch differierende Auffassungen deutlich werden (Eraut, 1998). Dies wird bereits bei der Verwendung grundlegender Begriffe deutlich, worauf Manzel (2007) auch im Hinblick auf die englischen Begriffe²⁰ hinweist, die häufig in deutschsprachigen Veröffentlichungen Verwendung finden:

„Besonders die englischen Begriffe *Competence* und *Competency* lassen sich im Deutschen leicht fehl interpretieren, denn der Begriff kann durch verschiedene andere Wörter beschrieben werden wie core skill, key qualification, fitness, ability, capacity, efficiency, proficiency. Der Terminus ‚Competence‘ bezieht sich hier sowohl auf die theoretische Konzeption als auch auf ein bestimmtes Kompetenzniveau einer speziellen Fähigkeit oder Fertigkeit, die eine Person erreicht hat, z.B. eine hohe Mathematikkompetenz. Der Begriff ‚Competency‘ wird verwendet, wenn es um eine besondere Kompetenzanforderung an eine Person geht, die diese erfüllt oder nicht, z.B. Teamfähigkeit.“ (S. 125f.)

Eine einheitliche Verwendung des Kompetenzbegriffs ist in der Forschung nicht zu erkennen (Frey, 2014, S. 733). In dieser Arbeit wird im Wesentlichen eine Definition von Kompetenzen verwendet, die sich an der bereits in Kapitel 1 dargestellten Formulierung von Klieme et al. (2007/2009) orientiert. Es werden hierbei die Charakteristika *Anwendungsbezug* (Kapitel 2.3.1.1), *Domänenspezifität* (Kapitel 2.3.1.2) sowie *Erlern- und Graduierbarkeit* (Kapitel 2.3.1.3) betont und in den Folgekapiteln näher erläutert.

2.3.1.1 Anwendungsbezug

Im Alltagsverständnis wird einer Person oftmals eine bestimmte Kompetenz zugesprochen, wenn sie eine Anforderungssituation in besonderer Qualität bewältigt, wobei der Grad der Bestimmtheit einer Kompetenz mit der Konkretion der Anforderungssituation steigt. Ob eine

²⁰ Für eine Übersicht über die historische Entwicklung des Kompetenzbegriffs im englischsprachigen Raum vgl. Eraut (1994/2000, S. 163ff.).

Person beispielsweise gut Auto fahren kann, wird im Alltag vermutlich weniger über das Vorhandensein des Führerscheins definiert, sondern dadurch, wie diese Person das Kraftfahrzeug im Straßenverkehr bewegt. Eine noch konkretere Aussage in diesem Alltagsverständnis wäre vielleicht noch möglich, wenn eine Person in einer besonders herausfordernden Situation, zum Beispiel einem drohenden Unfall, angemessen reagieren würde. Der Maßstab zur Beurteilung des Ausprägungsgrades einer Kompetenz wäre demnach die eigentliche Performanz im Sinne des Handelns der Person.

In einer solchen Vorstellung liegt Kompetenz gewissermaßen ‚hinter‘ der Performanz und wird daher auch als *Befähigung zur Handlung* bzw. *Fähigkeit* in Form einer „combination of attributes underlying some aspects of successful professional performance“ (Gonzi et al., 1993, S. 5f., zitiert nach Eraut, 1998, S. 134) bezeichnet – oder in einer Formulierung von Messick (1984, S. 227):

„Competence refers to what a person knows and can do under ideal circumstances, whereas performance refers to what is actually done under existing circumstances. Competence embraces the structure of knowledge and abilities, whereas performance subsumes as well the process of accessing and utilizing those structures and a host of affective, motivational, attentional and stylistic factors that influence the ultimate response.“ (Zitiert nach Eraut, 1998, S. 134)

Diese Erläuterung zum Verhältnis von Kompetenz und Performanz zeigt eine gewisse Nähe zur breiten Definition von Kompetenz bei Weinert (2001). Beide Auffassungen sind miteinander insofern vereinbar, als Kompetenz „attributes with performance“ (Eraut, 1998, S. 134) integriert. Das bedeutet, „a performance context rather than an education context is needed for such an integration to take place, thus emphasizing the closeness of competence to performance rather than its separation“ (ebd.). Auch bei Weinert (2001, S. 27f.) wird der Zusammenhang zwischen Fähigkeiten und Bereitschaften betont. Schließlich ist zum Handlungsbezug des in dieser Arbeit verwendeten Kompetenzbegriffs wichtig zu betonen, dass dieser zwar als komplexes Konstrukt angesehen wird, das nach Weinert (ebd.) eben auch motivationale und volitionale Aspekte beinhaltet, dabei jedoch gleichzeitig davon ausgegangen wird, dass es möglich ist, die einzelnen Kompetenzaspekte sowohl analytisch voneinander getrennt zu modellieren (Lindmeier, 2011, S. 34) als auch separat zu erheben, denn: „the pragmatic separation of cognitive components from further e.g. motivational aspects or interests is very common (cf. also Klieme & Leutner, 2006)“ (ebd., S. 35).

2.3.1.2 Domänenspezifität

In nahezu allen Kompetenzdefinitionen findet sich der Aspekt der Kontextspezifität (Klieme et al., 2007, S. 6) und es wird betont, dass, „[w]ährend in der Intelligenzforschung kognitive Leistungskonstrukte untersucht werden, die über eine breite Vielfalt von Situationen generalisierbar sind, ... sich Kompetenzkonstrukte auf spezifische Anforderungsbereiche [beziehen] – die Frage ‚*kompetent wofür?*‘ ist notwendiger Bestandteil jeder Kompetenzdefinition“ (ebd.). Eraut (1998) stellt die Kontextabhängigkeit von Kompetenzen in seiner Unterscheidung von Kompetenzen und Fähigkeiten in einem Modell dar (Abbildung 9).

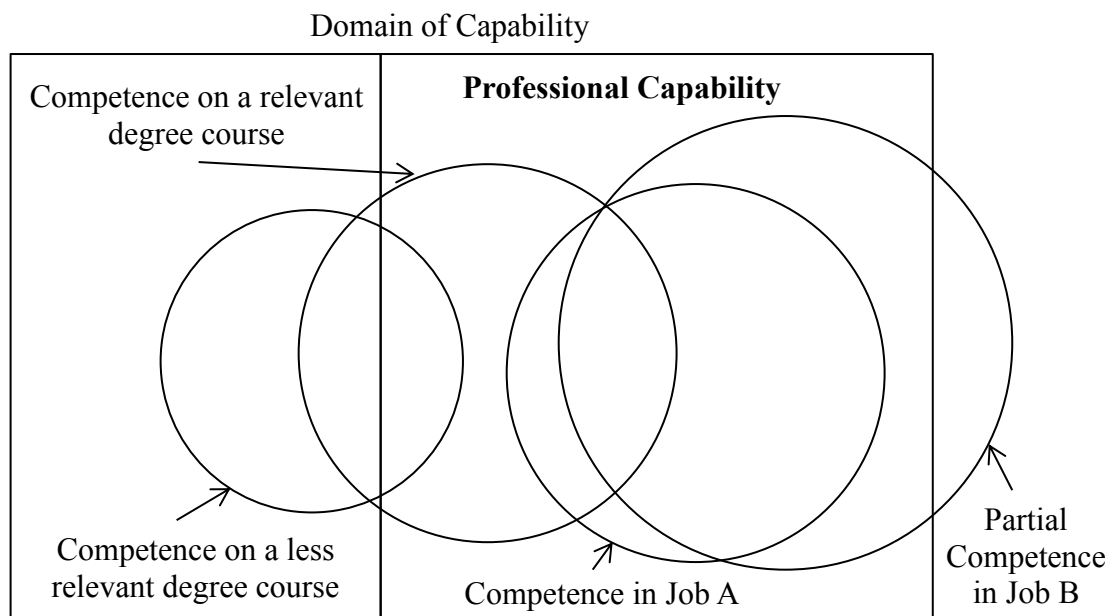


Abbildung 9: Zusammenhang von Fähigkeiten und Kompetenzen (eigene Darstellung nach Eraut, 1998, S. 135)

Entsprechend diesem Modell ist „competence ... socially situated and job referenced, capability is individually situated and profession referenced“ (ebd., S. 135). Obwohl dieser an der unmittelbaren Konkretion beruflicher bzw. professioneller Anforderungssituationen orientierte Referenzrahmen (capability-profession/competence-job) für die Definition von Kompetenzen und Fähigkeiten in den bislang vorgestellten Kompetenzdefinitionen nur bedingt

berücksichtigt ist, scheint er mit ihnen vereinbar, wenn näher beleuchtet wird, wie Eraut den Zusammenhang von Fähigkeiten und Kompetenzen beschreibt:

- „(1) Current competence is by definition part of a person's capability. Such competence is normally inferred from evidence of job performance.
- (2) The range of current competence is most easily extended when further learning can build on areas of additional capability. But this potential has to be recognised both by the professionals involved and by their employers.
- (3) Part of professional's capability involves being able to develop or transform one's practice over time, to create new knowledge through one's practice as well as learning from others.“ (Ebd., S. 136)

Bezogen auf den Lehrerberuf ist beispielsweise die Fähigkeit zur Kommunikation etwas „Alltägliches und Selbstverständliches“ (Hallet, 2006, S. 127). In unterschiedlichen schulischen Anforderungssituationen kann sie als kommunikative Kompetenz unverzichtbar sein, unter anderem im Rahmen institutioneller Diskurse (Dienstgespräch mit dem Rektorat, Kommunikation mit dem Schulministerium) oder als Fähigkeit zur Begleitung eines Lehr-Lern-Diskurses (gelenktes Unterrichtsgespräch, Moderation einer Pro-Kontra-Debatte im Politikunterricht) (ebd.). Nicht alle Facetten professioneller Kommunikationsfähigkeit werden sich in diesen unterschiedlichen Anforderungssituationen zeigen, sondern nur die jeweils notwendigen. Da Kompetenzen also in Abhängigkeit zur jeweiligen Anforderungs- oder Problemsituation stehen, bringt dies besondere Herausforderungen für die Forschung mit sich, da adäquate Messverfahren entwickelt werden müssen, die die spezifischen Anforderung des jeweiligen Kontextes konstant halten, denn *„context affects the development level or stage of a person's competence even when the effects of experience and domain are controlled for“* (Fischer, Bullock, Rotenberg & Raya 1993, S. 96).

Wie das Beispiel der kommunikativen Kompetenzen von Lehrkräften zeigt, muss der Begriff der Kontextspezifität noch genauer geklärt werden. Hierzu weisen Klieme et al. (2007) darauf hin, dass „[d]er relevante Kontext für die Definition eines Kompetenzkonstrukts ... einerseits hinreichend konkret sein [muss], ... andererseits auch nicht zu eng gefasst sein [sollte], da sonst einfaches Sachwissen oder isolierte Fertigkeiten unnötigerweise als Kompetenzen etikettiert werden“ (S. 8), und schlagen daher als Kriterium vor, dass „der Bezug auf eine Menge hinreichend *ähnlicher realer Situationen*, in denen bestimmte, ähnliche Anforderungen bewältigt werden müssen“ (ebd.) gegeben sein sollte. Dies ermöglicht dann, „dass mit Hilfe des Konzepts der Kompetenz eine mögliche Unterscheidung zwischen reinem Wis-

sen (z.B. über grammatikalische Regeln) und dessen Anwendung (z.B. in einer kommunikativen Situation) vorgenommen werden kann“ (ebd.).

Die Kontextspezifität der im Folgenden im Fokus stehenden Kompetenzen von Politiklehrkräften soll im Sinne psychologischer Expertiseforschung durch den Begriff *Domäne* ausgedrückt werden. Ist dabei eine solche „Domäne der Anforderungskomplex ..., dessen Anforderungen man durch Handlungen und die sie voraussetzenden Kompetenzen (Dispositionen) bewältigen kann“ (May, 2007, S. 72), kann diese präzisiert werden. Allerdings muss, sofern in dieser Arbeit von einer „Domäne Politik“ (Weißeno, 2008, S. 11) gesprochen wird, auch berücksichtigt werden, dass es weiterhin Auseinandersetzungen darüber gibt, wie eng oder weit diese gefasst wird und welche theoretischen Bezugspunkte für eine entsprechende Begründung herangezogen werden. Eine eindeutige Definition der Domäne Politik in Form einer Logik des (Schul-)Faches (Seifried & Ziegler, 2009, S. 85) wie im Fach Mathematik ist schon deshalb nicht möglich, weil es in den einzelnen Bundesländern – teilweise sogar innerhalb dieser – verschiedene Fächerbezeichnungen sowie differierende inhaltliche Gewichtungen der sozialwissenschaftlichen Bezugsdisziplinen gibt. Die Passung anhand fachbezogener Bereiche, wie beispielsweise für die Lesekompetenz der Schüler/-innen in der *PISA*-Studie vorgenommen (Klieme et al., 2007, S. 14), erscheint im Hinblick auf die sich zum Teil deutlich voneinander unterscheidenden wissenschaftlichen Positionen zum Referenzrahmen politischer Bildung ebenfalls nicht vollends befriedigend. Zwar können in der vorliegenden Arbeit die wissenschaftstheoretischen und bildungspolitischen Debatten zur Frage des inhaltlichen Kerns sozialwissenschaftlicher Bildung nicht ausführlich referiert werden,²¹ jedoch sei exemplarisch anhand der Diskussion über adäquate Konzeptmodelle zur Entwicklung von Schülerkompetenzmodellen (Autorengruppe Fachdidaktik, 2011; Weißeno et. al, 2010) darauf hingewiesen, dass die Fragen nach dem theoretischen Referenzrahmen einer Domäne weiterhin diskutabel erscheinen. Eine Antwort hierauf ist jedoch für die (Weiter-)Entwicklung eines entsprechenden domänenspezifischen Lehrerkompetenzmodells unabdingbar, vor allem, wenn ein tragfähiges Gesamtmodell entwickelt werden soll.

In dieser Arbeit wird die Modellierung der Kompetenzen von Lehrkräften der PKP-Studie als theoretischer Bezugspunkt angesehen und es wird bei der Konzeptualisierung dem

²¹ Für eine Übersicht über die Kontroverse um eine politisch-demokratische Bildung vgl. May (2007, S. 80ff.), zum Verhältnis von politischer und ökonomischer Bildung vgl. die Positionen von Detjen (2006) sowie Kaminiski (2006) und zu Perspektiven einer sozialwissenschaftlich-integrativen Bildung vgl. Hedtke (2006, 2011).

Ansatz gefolgt, Politik als Kern (Massing & Weißenö, 1995) zu betrachten, was auch eine getrennte Analyse und Erhebung unterschiedlicher Teilkompetenzen ermöglicht. Gleichzeitig ist darauf hinzuweisen, dass entsprechende Forschungen zu den weiteren Bezugsdisziplinen notwendig – und durch den Rückbezug auf die vorliegenden Konzeptualisierungen theoretisch und forschungspraktisch möglich – sind.

2.3.1.3 Erlern- und Gradulierbarkeit

In gängigen Kompetenzdefinitionen wird betont, dass Kompetenzen grundsätzlich erlern- und graduierbar sind. In „Abgrenzung zu kognitiven Grundfunktionen ..., die in wesentlich geringerem Maße erlernbar und trainierbar sind (Weinert, 2001)“ (Klieme et al., 2007, S. 7), wird davon ausgegangen, dass von einer Person mit ausreichenden grundlegenden kognitiven Fähigkeiten auch spezialisierte Kompetenzen gezielt ausgebildet werden können, beispielsweise durch „äußere Interventionen und institutionalisierte Bildungsprozesse“ (Fleischer et al., 2013, S. 6).

Durch die Grundannahme der Erlernbarkeit „unterscheidet sich dieses pädagogisch-psychologische Verständnis von biologischen oder anthropologischen Ansätzen wie beispielsweise von Lepenies (1971), nach denen es sich bei Kompetenzen um angeborene Fähigkeiten handelt“ (Blömeke, 2013b, S. 26). Es wird schließlich betont, dass Kompetenzen relativ zeitstabil sind, sodass es Personen mit bereichsspezifisch ausgeprägten Kompetenzen möglich ist, auch dauerhaft effektiv in entsprechenden Anforderungssituationen zu agieren (Klieme & Hartig, 2007, S. 14). Sofern eine Person also nicht wiederholt den Anforderungen einer professionellen Problemsituation entsprechen kann, wird davon ausgegangen, dass sich eine solche sporadisch gezeigte effektive oder überlegende Performanz nicht auf ein latentes Personenmerkmal zurückführen lässt.

2.3.2 Modellierung von Kompetenzen

Um die Aspekte, die zur Bewältigung professioneller Anforderungs- und Problemsituationen notwendig sind, beschreiben und analysieren zu können, bedarf es einer Vorstellung davon, wie Kompetenzen organisiert sind und in welcher Beziehung sie in ihren Dimensionen zueinander stehen. Hierzu werden *Kompetenzmodelle* entwickelt. *Kompetenzstrukturmodelle*

dienen dabei der Differenzierung verschiedener Teilkompetenzen (Dimensionen) und der Darstellung einer Binnenstruktur durch die Konzeptualisierung von Kompetenzdimensionen und gegebenenfalls (Sub-)Facetten. *Kompetenzniveaumodelle* stellen die Stufung dieser Dimensionen und Facetten sowie deren Ausprägungsgrade in den Vordergrund und *Kompetenzentwicklungsmodelle* fokussieren die Sequenzierung von Kompetenzen. Die in dieser Arbeit besonders interessierenden Kompetenzstrukturmodelle sollen ausführlicher formuliert folgende Aufgaben erfüllen (Fleischer et al., 2013):

„Kompetenzstrukturmodelle befassen sich mit der Frage, welche und wie viele verschiedene Teilkompetenzen (Dimensionen) einer bestimmten Kompetenz differenzierbar werden können. Diese Dimensionen repräsentieren inhaltlich unterscheidbare Aspekte, anhand derer sich Kompetenzunterschiede zwischen Personen beschreiben lassen. Die Binnenstruktur von Kompetenzen ist dabei vornehmlich durch die in einem bestimmten Bereich zu bewältigenden Anforderungen und die hierfür als notwendig erachteten kognitiven Prozesse bedingt (vgl. Hartig und Klieme 2006).“ (S. 8)

Kompetenzstrukturmodelle können ein- oder mehrdimensional sein (vgl. auch Abbildung 6). Mehrdimensionale Modelle werden vorgeschlagen, da deren „Erfassung mehrerer Teilkompetenzen mit einem mehrdimensionalen Messmodell ... die Prüfung der prognostizierten Kompetenzstruktur und eine differenziertere auf die Teilkompetenzen bezogene Diagnostik [erlaubt]“ (Mayer & Wellnitz, 2014, S. 25).

Bei der Modellierung der Binnenstruktur von Kompetenzen steht die Frage nach dem *Auflösungsgrad* im Vordergrund. Nach Weinert (2001) „kann eine Kompetenz in Teilkompetenzen im Sinne derjenigen Fähigkeiten und Fertigkeiten sowie Bereitschaften gegliedert werden, wie sie zur Lösung von Problemen in bestimmten Teilbereichen einer Domäne benötigt werden“ (K. Neumann, 2013, S. 36), wobei es neben theoretischen Grenzen der *Feingliedrigkeit* der Modellierung auch für die Testentwicklung relevante Fragen nach der *Feinkörnigkeit* der Messung zu diskutieren gilt. So weisen Klieme et al. (2007) darauf hin, dass „[b]ei der Entscheidung für ein bestimmtes Strukturmodell, also der Frage, wie differenziert spezifische Kompetenzen betrachtet werden sollen, ... jeweils eine Abwägung ökonomischer und theoretischer Aspekte vorgenommen werden [muss]“ (S. 13). Wenn „z. B. interindividuelle Unterschiede in verschiedenen Leistungsmaßen sehr hoch korreliert [sind], stellt sich die Frage, ob es nicht ökonomischer ist, diese Maße bei der Datenerhebung und der Auswertung zusammenzufassen und als eine gemeinsame Skala zu behandeln“ (Hartig & Klieme, 2006, S. 133). Schließlich wird betont, dass „[d]ie Validität des Modells und seiner Operationalisie-

runge ... zudem durch die Einbindung in geeignete kognitionspsychologische und fachdidaktische Analysen gesichert werden [sollte] (z. B. durch qualitative Analysen aus Fallstudien)“ (Leuders & Sodian, 2013, S. 32). Die Modellierung von Kompetenzen ist in der Regel der erste Schritt in Richtung der Nutzung von Kompetenzmodellen (Abbildung 10).

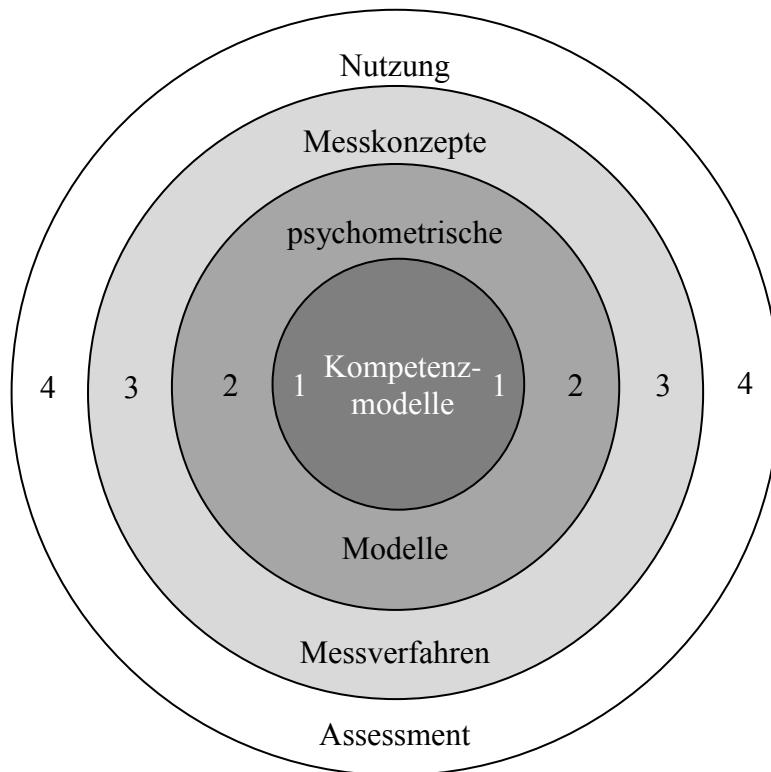


Abbildung 10: Darstellung der Entwicklungs- und Forschungsschritte von der Kompetenzmodellierung bis zur Kompetenzdiagnostik (eigene Darstellung in Anlehnung an Fleischer et al., 2013, S. 7)

Nach der Konzeptualisierung eines adäquaten Kompetenzmodells (1) folgt in der Regel die Entwicklung geeigneter psychometrischer Modelle (2). Anschließend werden Messverfahren erprobt (3), die es ermöglichen, die Kompetenzdimensionen und -facetten des Modells zu erfassen. Ziel ist schließlich die Nutzung des Kompetenzmodells, unter anderem in Form standardisierter Kompetenzdiagnostik (4) (Fleischer et al., 2013, S. 7f.).

2.3.3 Messung von Kompetenzen

Die Anlässe für die Beurteilung von Lehrpersonen – im Sinne der vorliegenden Arbeit also die Diagnostik von Lehrerkompetenzen – können vielfältig sein. So fassen Teddlie, Stringfield und Burdett (2013) die Bereiche „accountability, promotion, and staff development“ (S. 13) zusammen. Da Lehrer/-innen ein wesentlicher ‚Bestandteil‘ des Bildungssystems sind, wird „für die laufende Beobachtung der Qualität von Bildungssystemen (*system-monitoring*)“ (Hartig & Jude, 2007, S. 17) die Kompetenzdiagnostik als wichtig erachtet (ebd.).

In den vorhergehenden Kapiteln sind an zahlreichen Stellen bereits besondere Herausforderungen angeklungen, die mit der standardisierten Erhebung von Kompetenzen einhergehen können. In der folgenden kritischen Diskussion ausgewählter Ansätze der Kompetenzdiagnostik soll auch betrachtet werden, welche spezifischen Problemstellungen und Anforderungen sich bei kompetenzdiagnostischen Verfahren im Hinblick auf die üblichen Testgütekriterien (*Objektivität, Reliabilität und Validität*, vgl. auch Bortz & Döring, 2006, S. 193f.) identifizieren lassen. Nach der Übersicht über Grundlagen der Kompetenzdiagnostik (Kapitel 2.3.3.1) folgt in diesem Kapitel eine Darstellung von Herausforderungen, die sich im Zusammenhang mit der Konzeptualisierung des politikdidaktischen Wissens und Könnens in dieser Arbeit ergeben können und die in den Folgekapiteln bei der Modellierung, der Operationalisierung der Konstrukte sowie der Entwicklung geeigneter Testverfahren zu berücksichtigen sind. Hierzu werden Ansätze computerbasierter Erhebungsverfahren vorgestellt und diskutiert (Kapitel 2.3.3.2). In Kapitel 2.3.3.3 werden schließlich qualitative Kriterien zur Beurteilung von Antworten reflektiert, die in kompetenzdiagnostischen Verfahren zur Bewertung offener Antworten Anwendung finden müssten.

2.3.3.1 Grundlagen und Anforderungen domänenspezifischer Kompetenzdiagnostik

Wenn nach kognitionspsychologischem Verständnis Kompetenzen als „latente (kognitive) Fähigkeiten“ (Lehmann-Grube & Nickolaus, 2009, S. 63) modelliert werden, folgt daraus, dass sie nicht direkt beobachtbar und daher auch keiner direkten Erfassung (Messung) zugänglich sind (ebd.). Forschungsmethodisch sind dabei unterschiedliche Zugänge zur Erfassung von Kompetenzen denkbar. So unterscheiden Maag Merki und Werner (2014) die Methoden *Test*, *schriftliche Befragung (Fragebogen)*, *mündliche Befragung (Interview)*, *Doku-*

mentation und *Beobachtung* (S. 751ff.), wobei die besondere Bedeutung von Leistungstests für Kompetenzmessungen bei Lehrkräften hervorgehoben wird (ebd., S. 751). Riese und Reinhold (2014, S. 260) bezeichnen Leistungstests als einzig aussagekräftige Methode zur Kompetenzdiagnostik; diese steht daher in den nachfolgenden Überlegungen im Fokus.²²

Im Allgemeinen sind Leistungstest dadurch gekennzeichnet, dass ein Beurteilungsmaßstab für die Antworten vorliegt, zum Beispiel *richtig* und *falsch* (Bortz & Döring, 2006, S. 190; Maag Merki & Werner, 2014, S. 751). Während bei einfachen Richtig-falsch-Aufgaben oder Multiple-Choice-Fragen, bei denen nur eine Antwort richtig ist, von einer hohen *Auswertungsobjektivität* ausgegangen werden kann (Bortz & Döring, 2006, S. 195), kann diese Antworten auf offene Fragen aufgrund der Abhängigkeit der Beurteilung vom Auswertenden beeinträchtigt sein (Sedlmeier & Renkewitz, 2008, S. 71). Dies tangiert ebenso die *Interpretationsobjektivität*, als die „Frage, ob verschiedene Anwender aus demselben Testergebnis die gleichen Schlüsse ziehen“ (ebd.). Es bedarf daher bei offenen Antworten der Angabe von Normen im Sinne von Referenzpunkten, anhand derer eine Bewertung der Antworten erfolgen kann (Bortz & Döring, 2006, S. 195). Soll ein Test darüber hinaus mit externen Bezugsnormen verglichen werden, müssen diese auch vorhanden sein, was bereits in Kapitel 2.1 im Zusammenhang mit der Diskussion um einen leistungsbasierten Expertisebegriff für die Lehrerforschung problematisiert wurde (vgl. auch Besser, 2014, S. 16).

Entsprechend der Unterscheidung in *typisches Verhalten* und *maximale Anstrengung* (Schmidt-Atzert, 2012, S. 183) beziehen sich kompetenzdiagnostische Verfahrensweisen in Form von Leistungstests als „Instrumente zur Erfassung von Denkleistung“ (Wilhelm & Nickolaus, 2013, S. 24) auf maximale Anstrengung (ebd.). Hieran kann problematisiert werden, dass Messkonzepte, die das Kompetenzkonstrukt im Sinne maximaler Anstrengung verwenden, davon ausgehen, dass Versuchsteilnehmer/-innen auch gewillt sind, sich maximal anzustrengen (Schmidt-Atzert, 2012, S. 183). Individuelle Leistungsunterschiede können jedoch nur dann als Fähigkeitsunterschiede angesehen werden, wenn dies tatsächlich zutrifft (*Kompetenz-Performanz-Problem*). Daneben können weitere Einflussfaktoren die *Reliabilität* – also die „Zuverlässigkeit oder Messgenauigkeit eines Messinstrumentes“

²² Für eine Übersicht über weitere methodische Zugänge zur Erfassung von Kompetenzen vgl. Maag Merki & Werner (2014, S. 751ff.).

(Sedlmeier & Renkewitz, 2008, S. 72) – beeinträchtigen, so beispielsweise die Motivation oder auch die Müdigkeit der Probanden (ebd.).

Eraut (1998) gibt weiterhin Folgendes zu bedenken: „A person’s capacity to perform is affected by the total demands of the job, emotional as well as physical; and work overload has become almost normal for professionals in the public sector“ (S. 136). Wollen kompetenzdiagnostische Verfahren also auf berufliche Anforderungssituationen bezogene Kompetenzen zur Bewältigung von Problemen erfassen, müssten Testaufgaben dementsprechend realistische Kontexte widerspiegeln. So werden „[i]m Rahmen von Unterricht und seinem schulischen Umfeld ... Aufmerksamkeit, Anstrengung, Geduld und Ausdauer, Leistungsmotivation, Zielorientierung, Belohnungsaufschub und Selbstregulation, aber auch Emotionskontrolle und soziale Rücksichtnahme, Hilfsbereitschaft und Aushandlung von Interessen, Übernahme von Verantwortung, Kooperation oder konstruktive Konfliktbewältigung thematisch“ (Baumert & Kunter, 2006, S. 473), was nicht nur im Rahmen von Kompetenzmodellierung, sondern auch bei der Entwicklung adäquater Erhebungsinstrumente berücksichtigt werden sollte.

Die in Kapitel 2.3.1.2 angesprochene Domänenspezifität von Kompetenzen birgt dabei eine gewisse Problemstellung für deren Erhebung, da „[i]n *professionellen Domänen* ... Expertenwissen um Schlüsselkonzepte und eine begrenzte Anzahl von Ereignisschemata arrangiert [ist], an die Einzelfälle, episodische Einheiten oder Sequenzen von Episoden (Skripts) ange-dockt sind“ (ebd., S. 483). Infolgedessen sollte in kompetenzdiagnostischen Verfahren nach Eraut (1998) „current competence“ (S. 136) erfasst werden. Diese ist „normally inferred from evidence of job performance“ (ebd.), weshalb hier konkrete, hoch kontextualisierte Aufgabenstellungen erforderlich sind. Zwar kann vor dem Hintergrund der in dieser Arbeit verwendeten kompetenztheoretischen Grundlagen davon ausgegangen werden, dass (Teil-)Kompetenzen hoch aufgelöst modelliert und operationalisiert werden können (K. Neumann, 2013, S. 36) und „[d]ie theoretische Grenze der Feingliedrigkeit ... dann erreicht [ist], wenn zur Lösung eines Problems nur noch einzelne Fähigkeiten oder Fertigkeiten oder Bereitschaften benötigt werden“ (ebd.). Jedoch muss gleichzeitig bei der Erfassung isolierter Fähigkeiten bedacht werden, dass diese dann unter Umständen keine Teilkompetenz mehr im Sinne des zugrunde liegenden Kompetenzkonstrukts sind (ebd.). Im Hinblick auf die forschungspraktische Ausgestaltung wird daher auf die Wichtigkeit der Forschungsfrage sowie

des Kontextes der Erhebung hingewiesen (ebd.). Dementsprechend wäre der Informationsgehalt von Studien, die auf einem Kompetenzmodell mit niedrigem Auflösungsgrad beruhen und differenzierte Erkenntnisse zur Struktur von Teilkompetenzen oder interindividuellen Leistungsunterschieden von Probanden bzw. Probandengruppen liefern sollen, eher gering (ebd.).

Für die Entwicklung von Aufgabenstellungen und Test-Items zur Erfassung professioneller Kompetenzen ist im Zusammenhang mit der Festlegung des Auflösungsgrades eines Kompetenzmodells weiterhin Folgendes zu berücksichtigen: „demands, tasks, and problems ... resemble each other sufficiently with respect to the terms of reference“ (Lindmeier, 2011, S. 36). Hiermit ist auch die Frage nach der *Validität* von Leistungstests angesprochen. Dieses Gütekriterium „eines Tests gibt an, wie gut der Test in der Lage ist, genau zu messen, was er zu messen vorgibt“ (Bortz & Döring, 2006, S. 200). Unterschieden werden die *Inhaltsvalidität*, die *Konstruktvalidität* und die *Kriteriumsvalidität* (ebd., S. 200ff.). Die *Inhaltsvalidität* eines Leistungstests kann dabei als gegeben angesehen werden, wenn „der Test das zu messende Merkmal in seinen wesentlichen Aspekten erschöpfend erfasst“ (Sedlmeier & Renkewitz, 2008, S. 77). Je mehr Test-Items aus der Menge derjenigen Items, die für eine Operationalisierung der Konstrukte infrage kommen, enthalten sind, desto höher ist die Inhaltsvalidität (Bortz & Döring, 2006, S. 200). Die Testaufgaben müssen dem Referenzrahmen der modellierten Kompetenzfacetten für eine differenzierte Diagnostik also entsprechen, was sich vor dem Hintergrund der thematisierten Theorie- und Forschungsdesiderata für die Politikdidaktik als schwierig erweisen kann, denn was eine *typische* Unterrichtssituation im jeweiligen Fach ist, anhand derer Items formuliert werden, muss zunächst politikdidaktisch geklärt werden. Wie bereits anhand des Beispiel-Items (Abbildung 5) aus der PKP-Studie verdeutlicht, wäre der Referenzrahmen der modellierten Kompetenzfacetten idealerweise normativ-empirischer Art, was jedoch das Vorhandensein adäquater Forschungen notwendig macht. Sofern Testaufgaben den Referenzrahmen der Kompetenzmodellierung inadäquat widerspiegeln (Lindmeier, 2011, S. 37) oder der Referenzrahmen selbst – wie in den Überlegungen zur Kompetenzstruktur von Politiklehrkräften in den bislang vorliegenden bildungsadministrativen Vorgaben zur Lehrerbildung bereits thematisiert – unklar ist, kann nicht gewährleistet werden, dass die jeweils interessierenden Kompetenzaspekte tatsächlich erfasst werden (ebd.).

Bei der Interpretation der Testleistungen werden die Ergebnisse von Lehrkräften teilweise in Verbindung gesetzt mit Kriterien, „mit denen die Testwerte zusammenhängen sollten, sofern der Test tatsächlich das misst, was er zu messen vorgibt“ (Sedlmeier & Renkewitz, 2008, S. 78). Entsprechende Zusammenhangsmaße müssen sich dabei an der Güte der *Kriteriumsvalidität*, die „die Übereinstimmung zwischen den Testwerten und so genannten Kriterien bestimmt“ (Sedlmeier & Renkewitz, 2008, S. 78), messen lassen. Problematisch ist hierbei, dass oftmals keine Außenkriterien gefunden werden können (ebd.; vgl. auch Bortz & Döring, 2006, S. 201), die indikativ für Lehrerkompetenzen sind und ihrerseits auch wiederum den üblichen Testgütekriterien genügen müssen. Da sich das professionelle Wissen und Können von Lehrkräften nach einem funktionalen Verständnis auf die Aufgabe des Unterrichtens bezieht (vgl. auch Kapitel 2.1 und Kapitel 2.3.3.3), läge es nahe, im Sinne prognostischer Validität (Bortz & Döring, 2006, S. 200f.) die Testleistungen mit dem Kompetenzzuwachs bei den Lernenden in Zusammenhang zu setzen. Hieran wird jedoch kritisiert, „dass die Lernleistungen von Schüler/innen durch verschiedene Kontextparameter beeinflusst wird und folglich ein Rückschluss auf Lehrkompetenz mit einem ‚black box – Irrtum‘ gleichzusetzen ist“ (Maag Merki & Werner, 2014, S. 751; vgl. auch Lindmeier, 2014a, S. 67).

Eine besondere Herausforderung für die Entwicklung von Testverfahren zur Kompetenzdiagnostik ergibt sich schließlich, wenn Kompetenz – wie in der vorliegenden Arbeit – als Einheit von Wissen und Können aufgefasst wird. Zwar ist „[e]s nicht unüblich, ‚Können‘ über latente Variablenmodelle zu konzeptionieren und diese reduktionistisch anzulegen“ (Alisch, Hermkes & Möbius, 2009, S. 251). Jedoch „wäre kritisch zu erörtern, inwiefern derzeit gebräuchliche Kompetenzoperationalisierungen tatsächlich vorrangig deklaratives und prozedurales Wissen aus einem jeweils umschriebenen Gegenstandsbereich erfassen“ (Wilhelm & Nickolaus, 2013, S. 25), denn, so Frey und Hartig (2013):

„[Es] rücken zunehmend Kompetenzen in den Blick, die im teilweise zeitkritischen Ausführen komplexer Interaktionen bestehen und/oder die an realistische Anwendungssituationen gebunden sind. Die Quantifizierung von Interaktionsmustern, die hochinferente Protokollierung von Bearbeitungszeiten und die Darstellung realistischer Kontexte ist mit dem statischen Testmedium Papier- und Bleistift in der Regel nur sehr begrenzt oder mit sehr großem Aufwand möglich. Computer bieten hier vielfältigere Möglichkeiten, da in computerbasierten, interaktiven Testumgebungen Daten verhaltens- und zeitkritisch erhoben werden.“ (S. 54)

Auch Oser et al. (2010) kritisieren, dass „die meisten Diagnoseinstrumente Kompetenz unabhängig vom situativen Kontext zu erfassen versuchen“ (S. 6). Da in dieser Arbeit der Auffassung von Kompetenzen nach Klieme und Hartig (2007) als anforderungsspezifische Leistungsdispositionen gefolgt wird, die „weiter gefasst sind als die o. g. Wissens- und Wahrnehmungskonstrukte“ (Lindmeier, 2013, S. 45), stellt sich die Frage nach den Möglichkeiten, diesen Anforderungsreichtum des Lehrerberufs adäquat in Erhebungsverfahren abzubilden.

Klassische schriftliche Tests scheinen hierfür ungeeignet. Unter Rückbezug auf das Modell zu Konzepten des Lehrerwissens nach Neuweg (2014, S. 585) kann festgestellt werden, dass in schriftlichen kompetenzdiagnostischen Verfahren „die Erfassung von Wissen III ausgeschlossen [ist], während Wissen I zweifellos erfasst werden kann“ (Vogelsang & Reinhold, 2013, S. 107). Das *Handeln-Können* der Teilnehmenden wird insofern nicht valide erfassbar sein. Da zur Bearbeitung von Aufgaben in einem Fragebogen zunächst eine (interne) Kommunikation stattfinden muss, um schließlich die Gedanken external (schriftlich) verbalisieren zu können, ist es notwendig, Wissen zu explizieren. Nach der in Kapitel 2.2 erläuterten Transformationsposition würde in einem solchen Testverfahren daher zwar davon ausgegangen werden, dass explizites (deklaratives) Wissen 2 (Neuweg, 2014, S. 584) erfassbar wäre, nicht jedoch implizites Wissen. Den Theorien zum impliziten Wissen zufolge würde „[e]in direkter Wirkmechanismus von Wissen I auf Wissen II ... negiert, so dass das in einem Test erfasste Wissen nicht direkt als handlungsvalide bezeichnet werden kann“ (Vogelsang & Reinhold, 2013, S. 113). Insofern liegt eine Herausforderung für die Forschung auf methodischer Seite darin, Instrumente zur Erhebung handlungsnahen Wissens zu entwickeln (Lehmann-Grube & Nickolaus, 2009, S. 65). Es wird „[m]it Hilfe der videobasierten Methoden ... versucht, diese Kritikpunkte aufzugreifen und die professionellen Anforderungen des Unterrichtens stärker zu berücksichtigen“ (Lindmeier, 2013, S. 52), weil solchen Verfahren „das Potential zugeschrieben [wird], im Vergleich zu schriftlichen Maßen eine höhere Validität erreichen zu können“ (ebd.). Hiermit ist in diesem Zusammenhang die *Konstruktvalidität* der modellierten handlungsnahen Aspekte professioneller Kompetenzen angesprochen, die Sedlmeier und Renkewitz (2008) folgendermaßen zusammenfassen:

„Die Konstruktvalidierung eines Tests ist ein längerer, fortdauernder Prozess, in dem theoretische Aussagen über das zu messende Merkmal mit Hilfe des Tests überprüft werden. Die Validität des Tests wird hier also nicht nur anhand einzelner Außenkriterien ermittelt, son-

dern indem geprüft wird, ob möglichst vielfältige Hypothesen über das Merkmal durch die Testwerte bestätigt werden. Dies setzt natürlich voraus, dass solche Hypothesen abgeleitet werden können.“ (S. 79)

Ein Teilaspekt der Konstruktvalidität ist die *diskriminante Validität*, die „fordert, dass sich das Zielkonstrukt von anderen Konstrukten unterscheidet“ (Bortz & Döring, 2006, S. 203), wobei Korrelationen als Indikatoren für die diskriminante Validität eines Tests angesehen werden können (Krauss, et al., 2011, S. 150; Weschenfelder, 2014a, S. 190). Je nachdem, wie hoch die operationalisierten Konstrukte miteinander korrelieren, kann dies als Hinweis dafür angesehen werden, dass die entsprechenden Skalen das Gleiche oder etwas anderes messen (vgl. auch Kapitel 2.2 zum Zusammenhang der Konstrukte des fachdidaktischen Wissens und des Fachwissens).

Es lässt sich zusammenfassen, dass im Hinblick auf den Einsatz von schriftlichen Tests kritisch reflektiert werden muss, ob diese eine valide Erhebung handlungsnaher Aspekte professioneller Kompetenzen ermöglichen können. Selbst wenn im Sinne einer Transformationsposition davon ausgegangen wird, dass die häufig in Form von Wissenstests erfassten kognitiven Fähigkeiten und Fertigkeiten von Lehrkräften eine Handlungsdisposition darstellen, besteht weiterhin die Herausforderung, Fragen zu konstruieren, die die Kontextspezifität und Authentizität unterrichtlicher Anforderungssituationen adäquat abbilden. Inwiefern videobasierte Verfahren der Kompetenzdiagnostik bei der Bewältigung dieser Herausforderungen für forschungsmethodische Zugänge unterstützen können, soll im folgenden Kapitel erörtert werden.

2.3.3.2 Möglichkeiten und Grenzen computerbasierter videogestützter Kompetenzdiagnostik

Für die empirische Unterrichtsforschung sind unterschiedliche Vorgehensweisen denkbar, Videografie als Forschungsmethode einzusetzen (für eine Übersicht vgl. Blömeke, 2013b, S. 35ff.), wobei sich das Format des *video-cued testing* als besonders zielführend für die adäquate Erfassung von Kompetenzen erweisen kann (ebd., S. 36). In diesem Testformat werden den Probanden kurze Videoausschnitte (*Vignetten*) als Stimulus gezeigt, auf die sie reagieren müssen. Die Rückmeldung kann auf unterschiedliche Weise erfolgen, zum Beispiel in Form von Likertskalen zur Beurteilung von Unterricht (ebd., S. 37) oder durch offene bzw. ge-

schlossene schriftliche oder mündliche Antwortformate zu möglichen Reaktionen auf Fehler oder Verständnisschwierigkeiten der Schüler/-innen.

Die Vorteile solcher videobasierten Verfahren sind vor allem, dass sie unterrichtsrelevante Faktoren wie *Anforderungsreichtum*, *Komplexität*, *Situativität* und – insbesondere wenn die Vignetten aus Videoaufnahmen von ‚echtem‘ Unterricht stammen oder auf diesem basieren – höhere *Authentizität* angemessener implementieren können, als es in traditionellen Testverfahren möglich ist. Werden zudem nicht nur die Stimuli videobasiert an die Probanden ausgegeben, sondern auch noch deren Antworten technisch erfasst, ist es möglich, *Spontaneität* (durch Zeitdruck) und *Unmittelbarkeit* (geringe Latenz zwischen videobasiertem Stimulus und Antwort der Probanden) mithilfe videobasierter Maße abzubilden (Knievel et al., 2015; Lindmeier, 2011, Lindmeier et al., 2012). Im Hinblick auf die spezifischen Anforderungen valider Kompetenzmessungen wird die Eignung Videovignetten-basierter Leistungstest zudem hervorgehoben, da sie proximale, also nahe Indikatoren²³ von Kompetenz objektiver erfassen können, als es beispielsweise in Verfahren der Selbsteinschätzung möglich ist (Rehm & Bölsterli, 2014, S. 215).

Zusätzlich zu den Vorteilen videobasierter Kompetenzdiagnostik sind mögliche theoretische sowie forschungspraktische Probleme zu diskutieren. So muss hinsichtlich der Bewertung von offenen Antworten der Probanden geklärt sein, welche Expertise die Bewertenden (*Rater*) selbst für eine valide Diagnostik haben müssen (*Samplingproblem*) und ob die in den *rating guidelines* festgelegten Bewertungskriterien der Komplexität des Ausgangsmaterials (Stimulus) sowie der Antworten gerecht werden (Blömeke, 2013b, S. 38f.). Zur Lösung des Samplingproblems wird von einigen Autoren das Rating durch externe Beobachter (*expert scoring*) als – vor allem im Vergleich zu Ratings durch Lehrer/-innen oder Schüler/-innen – besonders zielführend erachtet (Helmke, 2014, S. 292ff.). Es wird dabei angenommen, „dass Beobachter durch das absolvierte Training über eine hinreichende Beobachtungs- und Urteilskompetenz verfügen“ (Praetorius, 2014, S. 36). Jedoch muss neben den weiteren üblichen Quellen für einen Rater-Bias, der die Güte der Qualität eines Ratings maßgeblich beeinflussen kann (für eine Übersicht vgl. ebd., S. 37ff.), für die Bewertung der Qualität von Unterrichtsaspekten das Problem beachtet werden, dass hierbei „[u]nterschiedliche Item-

²³ Demgegenüber sind „[d]istale, also vom eigentlichen Unterrichtsgeschehen entfernte Indikatoren für die professionelle Kompetenz ... z.B. Ausbildungsstand, Zertifikate und Noten“ (Brovelli, Bölsterli, Rehm & Wilhelm, 2013, S. 307).

Interpretationen von Ratern ... als eine der Hauptursachen für Rater-Bias angesehen [werden] (z. B. Hoyt, 2000; Myford & Wolfe, 2003; Pietsch & Tosana, 2008)“ (ebd., S. 43). Gerade wenn die Problemstellung im Ausgangsmaterial für die Aufgabenformulierung uneindeutig ist, sind entsprechend komplexe Antworten zu erwarten, was hohe Anforderungen an das Ratingmanual stellt, sodass es „unerlässlich ist ..., wenigstens stichprobenweise die Übereinstimmung zwischen mehreren Beurteilern, die dieselbe Person beurteilen, zu untersuchen“ (Hartig & Jude, 2007, S. 20), um so eine ausreichende Objektivität der Beurteilungen sicherzustellen.

Die Videostimuli sollten vor allem im Hinblick auf die mit den üblichen Testgütekriterien verbundenen Ansätze zur Erhöhung des Standardisierungsgrades (Lindmeier, 2013, S. 53ff.) repräsentativ sein und zum Beispiel „besonders häufige Unterrichtssituationen [zeigen] oder auch solche, die für Schülerfehler zentral sind“ (Blömeke, 2013b, S. 39). Ein Kriterium, das das Typische an Lehr-Lern-Situationen bei der Sichtung des Ausgangsmaterials für die Gestaltung der Videovignetten fokussieren soll, erfordert jedoch, dass hierzu entsprechende Forschungserkenntnisse vorliegen. Wie bereits an anderen Stellen in dieser Arbeit angesprochen, stehen politikdidaktische Studien, die dieses Kriterium für die Erhöhung des Standardisierungsgrades von Videografieforschung konstituieren könnten, zurzeit teilweise noch aus.

Im Zusammenhang mit der Qualität der Videovignetten als Ausgangsmaterial für das video-cued testing sind weitere Herausforderungen zu berücksichtigen. So weisen Gronostay, Neumann und Manzel (2015, S. 161) auf das mögliche Problem von Kameraeffekten hin, die Auswirkungen auf die Authentizität der Unterrichtsaufnahmen haben können (vgl. auch Petko, Waldis, Pauli & Reusser, 2003, S. 270; Fankhauser, 2013). Zwar kommt der Authentizität der Unterrichtsaufnahmen in kompetenzdiagnostischen Verfahren eine insgesamt weniger große Bedeutung zu als beispielsweise in Forschungen, in denen der videografierte Unterricht selbst beurteilt wird, da die Unterrichtsausschnitte in der Kompetenzdiagnostik lediglich als Stimulus für die eigentliche Aufgabenstellung dienen und daher im Wesentlichen nur die Reaktion der Probanden interessiert. Dennoch dürfte die Qualität des Ausgangsmaterials insofern wichtig sein, als „der videografierte Unterricht als *kontext-sensitiver Impuls für die Datengewinnung* genutzt [wird]“ (Blömeke, 2013b, S. 36), weil die Realitätsnähe des Kontextes also im Hinblick auf die eingangs dargestellte Kritik an bisherigen traditionellen

Erhebungsverfahren in videobasierten Messungen besonders zu berücksichtigen ist. Für die Sicherstellung adäquater Validität ist „[t]he perceived authenticity of the situations depicted in the items regarding action-related competencies ... an important factor to ensure content validity“ (Lindmeier, 2011, S. 174), sodass bei der Verwendung nichtauthentischer Unterrichtsmaterialien (zum Beispiel in Form nachgespielter Unterrichtssituationen) Vorkehrungen zur Kontrolle der wahrgenommenen Authentizität angewendet werden sollten (ebd., S. 174f.).

Weitere Maßnahmen zur Standardisierung können sich hierbei auf die Videovignetten selbst beziehen, aber auch auf die Formatierung der Testaufgaben (Lindmeier, 2013, S. 53). Bei Lindmeier (2011) kommen beispielsweise zwar im Prinzip offene Aufgabentypen zum Einsatz, die darauf abzielen, dass die Testperson in der Rolle der Lehrkraft Lernenden in fachlichen Problemsituationen hilft, jedoch wird durch das aufgabenspezifische *framing* der Unterrichtssituation sowie durch abschließende Antwortvorgaben das mögliche Antwortmuster vorgegeben und somit in diesem Aspekt standardisiert. Darüber hinaus werden die Antworten anhand eines Kodierschlüssels (Experten-Rating anhand taxonomischer Kategorien) ausgewertet, um die Daten auch mittels quantitativer Verfahren analysieren zu können; daher findet auch hier eine Erhöhung des Standardisierungsgrades der Beurteilungen der Antworten statt (Lindmeier, 2013, S. 53). Dies sollte bei videobasierten Verfahren außerdem aus forschungsökonomischen Gründen – beispielsweise beim Einsatz von *Largescale*-Studien – erwogen werden, weil zusätzliche Schritte bei der Vorbereitung und Durchführung der Datenerhebung (Aufnahme von Unterrichtssituationen, Schnitt von Videovignetten, Vorbereitung der Computerarbeitsplätze) sowie bei der Datenanalyse (Transkription der Videodateien, Rating durch unterschiedliche Experten) zu erwarten sind. Schließlich ist zu berücksichtigen, dass Videografie als methodischer Zugang zur fachspezifischen Unterrichtsforschung in der politisch-sozialwissenschaftlichen Fächerdomäne weiterhin „ein Novum dar[stellt]“ (Manzel & Gronostay, 2013, S. 199) und, „[o]bwohl es bereits in den 90er Jahren erste Ansätze der Unterrichtsbeobachtung mit Videomitschnitten auch im Politikunterricht gegeben hat (Grammes & Weißeno, 1993; Schelle, 1996), ... bislang ein systematischer Zugang [fehlt]“ (ebd.).

Wenn videobasierte Erhebungsverfahren zur Anwendung kommen sollen, erfordert dies in der Regel eine computerbasierte Testumgebung. Wie bereits zu Beginn des Kapitels festge-

stellt, ist „[e]ine ganz wesentliche Stärke computerbasierter Testmethoden ... gerade für die Erfassung von Kompetenzen von besonderer Bedeutung: Die Testvorgabe am Computer erlaubt die Verwendung neuer und innovativer Testmaterialien, die bei gedruckten Medien nicht möglich ist“ (Jurecka & Hartig, 2007, S. 45). Unter *computerbasiertem Assessment* (*computer-based assessment*) – bzw. dem synonym verwendeten Begriff *computerbasiertes Testen* (*computer-based testing*) – wird eine Testumgebung verstanden, in der „zumindest die Vorgabe des Testmaterials am Computerbildschirm (on-screen presentation) [erfolgt], und die getestete Person ... ebenfalls über den Computer (z.B. Tastatur oder Maus) [reagiert]“ (ebd., S. 38) und „[d]ie einzelnen Reaktionen ... zumindest elektronisch aufgezeichnet [werden]“ (ebd.). Computerbasierten Erhebungsverfahren wird das Potenzial zugeschrieben, die klassischen Testgütekriterien Objektivität, Reliabilität und Validität durch die Möglichkeiten zur Erhöhung des Standardisierungsgrades zu verbessern (ebd., S. 44f.). Ebenso können sich testökonomische Vorteile ergeben (ebd., S. 45), da einmalig eingerichtete computerbasierte Erhebungen beliebig häufig und im Prinzip ohne Folgekosten der Sachmittel (wie beispielsweise Kopien von Fragebögen) eingesetzt werden können.

Diskutiert werden neben diesen Vorzügen auch kritische Aspekte wie eine mögliche *Äquivalenzproblematik* (*cross-mode equivalence*), was bedeutet, dass „[d]urch einen Wechsel des Testmodus, z.B. die Übertragung von Tests von einer Papier-Bleistift-Version in eine computerbasierte Version ...[,] sich unter Umständen die Natur des erfassten Konstrukts verändern [kann], wenn z.B. das Testverhalten durch die computerisierte Darbietung verändert wird“ (ebd., S. 41). Weiterhin kann der Einsatz von computerbasierten Erhebungsverfahren „die *Testfairness* zuungunsten von Personen beeinträchtigen ..., die im Umgang mit dem Computer weniger vertraut sind“ (ebd., S. 43). Schließlich müssen die technischen Voraussetzungen bedacht werden (ebd.). Wenn alle notwendigen technischen bzw. technologischen Ressourcen für ein computerbasiertes Verfahren vorhanden sind, dürften die bereits angesprochenen testökonomischen Vorteile überwiegen (ebd., S. 44f.). Sofern jedoch eine Videoaufnahme der Reaktion der Testperson erfolgen soll, ist auch hierzu entsprechendes technisches Equipment bereitzustellen (vgl. Kapitel 6.2 für weitere Details zu den technischen Aspekten der vorliegenden Studie).

Insgesamt lassen sich Weiterentwicklungsmöglichkeiten im Bereich der Kompetenzdiagnostik durch videobasierte Forschungsmethoden ausmachen (vgl. auch Weschenfelder, 2014a, S.

173f.). Ein höherer Standardisierungsgrad erfordert dabei in der Regel die Implementation des Tests in eine computerbasierte Umgebung, wodurch der technische Aufwand steigt. Als Leistungstest kann besonders das Format des video-cued testing eine Ausgangsbasis domänenspezifischer Kompetenzdiagnostik sein. Hierzu werden Videostimuli eingesetzt, auf die Probanden reagieren müssen, zum Beispiel in natürlicher Sprache in einem offenen Antwortformat. Zur Sicherung einer möglichst hohen Interpretationsobjektivität bei offenen Antworten sollten diese durch mehrere Rater beurteilt werden. Gleichzeitig ist hiermit angesprochen, dass es qualitativer Kriterien bedarf, um entsprechende Beurteilungen vornehmen zu können, worauf im folgenden Kapitel näher eingegangen werden soll.

2.3.3.3 Qualitative Merkmale politikdidaktischer Kompetenzen

Da „Kompetenzprofile ... zwar [definieren], was eine Lehrperson zur erfolgreichen Bewältigung des Berufsalltags können muss, sie ... aber nichts über die Güte aus[sagen]“ (Oser et al., 2010, S. 10), sind domänenspezifische Überlegungen zur qualitativen Beurteilung von Lehrerkompetenzen notwendig. Je komplexer die Reaktion einer Testperson ausfällt, desto mehr werden kriteriale Referenzpunkte benötigt, um eine entsprechende Beurteilung auch intersubjektiv vergleichbar zu machen. Dieser Umstand trifft insbesondere auf offene Antworten in kompetenzdiagnostischen Verfahren wie dem video-cued testing zu. Im Folgenden sollen daher Aspekte einer hohen Expertise fokussiert werden, die sich als Bezugsrahmen für die qualitative Beurteilung der politikdidaktischen Kompetenz von Politiklehrkräften in den Facetten des Wissens über Instruktions- und Vermittlungsstrategien sowie des Wissens über (Fehl-)Vorstellungen von Lernenden anbieten. Hierbei steht zwar die Lehrperson im Mittelpunkt der Betrachtung, aber in Anlehnung an die von Helmke (2014, S. 24ff.) charakterisierten, sich gegenseitig ergänzenden Qualitäten eines variablen- und eines personenzentrierten Ansatzes wird gleichzeitig der funktionale Aspekt der identifizierten Merkmale betont, die diese für eine hohe Unterrichtsqualität haben (vgl. auch Abbildung 1).

Es werden solche Kriterien fokussiert, die sich auf die in dieser Arbeit interessierenden Merkmale ‚guten Politikunterrichts‘ beziehen. Daher können nicht alle Dimensionen von Unterrichtsqualität angesprochen werden, sondern ausschließlich diejenigen, die mit den beiden zentralen politikdidaktischen Kompetenzfacetten in Zusammenhang stehen. Fachunabhängige Merkmale hoher Unterrichtsqualität wie *Klassenführung* und *Umgang mit*

Heterogenität werden infolgedessen zunächst nicht weiter betrachtet. Außerdem ist zu berücksichtigen, dass es in der Politikdidaktik bislang kaum gesicherte Vorstellungen dazu gibt, was ‚guter Politikunterricht‘ ist, also welche fachspezifischen Merkmale neben den etablierten fächerübergreifenden Merkmalen (Helmke, 2014) identifiziert werden können (Manzel & Gronostay, 2013, S. 200). Es ist daher notwendig, dass fächerübergreifende Merkmale herausgearbeitet werden, um auf dieser Grundlage mögliche domänenspezifische Implikationen reflektieren zu können.

Ein Modell, das in der Darstellung der vermuteten Zusammenhänge unterrichtlicher Wirkungen *Dimensionen guten Unterrichts* fokussiert, bieten Klieme, Lipowsky, Rakoczy & Ratzka (2006) an (Abbildung 11).

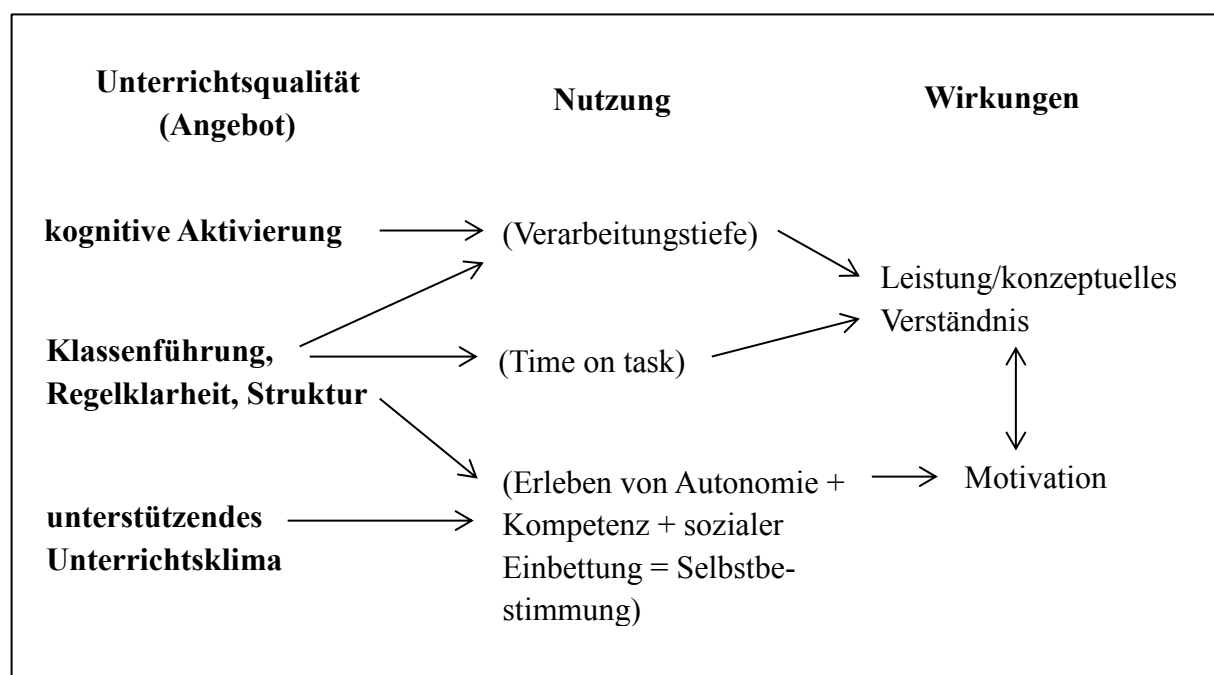


Abbildung 11: Grunddimensionen der Unterrichtsqualität und deren vermutliche Wirkung (eigene Darstellung in Anlehnung an Klieme, Lipowsky, Rakoczy & Ratzka, 2006, S. 131; Klieme & Rakoczy, 2008, S. 228)

Ähnlich wie beim Wirkmodell von Helmke (2014, S. 71) steht im Zentrum dieses Modells die Annahme, dass die Qualität des Angebots (Unterricht) beeinflusst, ob und wie die Schüler/-innen das Lernangebot wahrnehmen. Die im Modell abgezeichneten Merkmale hoher

Unterrichtsqualität werden mit Bezug auf die TIMSS-Videostudie (für eine Übersicht vgl. BMBF, 2001) begründet; in dieser wird „ein Vorschlag zur Formulierung von ‚drei Basisdimensionen guten Unterrichts‘ formuliert, der die empirischen Befunde jener Studie zum Zusammenhang zwischen Unterrichtsmerkmalen und Lernzuwächsen erklärte (Klieme/Schümer/Knoll 2001)“ (Klieme & Rakoczy, 2008, S. 228). Demnach können folgende Merkmale guten Unterrichts identifiziert werden:

- „strukturierte, klare und störungspräventive Unterrichtsführung,
- unterstützendes, schülerorientiertes Sozialklima,
- kognitive Aktivierung, zu der je nach fachlichem Kontext z.B. die Qualität der Nutzung von Experimenten im Physikunterricht, herausfordernde, offene Aufgaben in der Mathematik und generell ein diskursiver Umgang mit Fehlern gehören kann.“ (Ebd.)

Die Formulierungen zur Basisdimension der kognitiven Aktivierung deuten bereits an, dass diese jeweils fachspezifisch ausdifferenziert werden müssen, um sie einer Erhebung zugänglich machen zu können (ebd., S. 229). Im COACTIV-Forschungsprogramm werden diese drei Dimensionen hoher Unterrichtsqualität ebenfalls als Referenzpunkte genannt. Als qualitätsrelevante Merkmale des Unterrichts, die von den Lehrkräften beeinflussbar sind, werden unter Rekurs auf das Angebots-Nutzungs-Modell unterrichtlicher Wirkungen (Helmke, 2014, S. 71; vgl. auch Abbildung 1) folgende Dimensionen identifiziert: „(1) kognitiv herausfordernde Lerngelegenheiten in den Kernbereichen der jeweiligen Domäne, (2) Lernunterstützung durch sorgfältige Überwachung des Lernprozesses, individuelle Rückmeldung und adaptives Unterrichten und (3) ein effizienteres Klassen- und Zeitmanagement des Unterrichtsverlaufs“ (Baumert et al., 2011, S. 13), wobei in COACTIV domänenspezifisch das kognitiv anregende Potenzial von mathematischen Aufgaben, die individuelle Unterstützung durch die Lehrkraft sowie das Classroom Management (ebd., S. 13f.) als diejenigen Dimensionen abgeleitet wurden, die „die Grundlage für das Modell der Unterrichtsqualität, das in COACTIV als ein Kriterium für erfolgreiches Handeln der Lehrkräfte herangezogen wird“ (S. 14), darstellen.

In einer umfangreichen fächergenerellen Übersicht fasst Helmke (2014) diese Qualitätsbereiche zusammen:

1. Klassenführung
2. Klarheit und Strukturiertheit
3. Konsolidierung und Sicherung
4. Aktivierung

5. Motivierung
6. lernförderliches Klima
7. Schülerorientierung
8. Kompetenzorientierung
9. Umgang mit Heterogenität
10. Angebotsvariation“ (S. 168f.)

Von diesen Merkmalen scheinen für die weiteren domänenspezifischen Überlegungen zu qualitativen Aspekten der Expertise, die sich formal auf guten Unterricht beziehen, insbesondere die Qualitätsbereiche *Klarheit und Strukturiertheit, Aktivierung, Motivierung, Schülerorientierung, Angebotsvariation* sowie *Kompetenzorientierung* relevant zu sein, da in ihnen in den Ausformulierungen bei Helmke ein stärkerer Fachbezug hervorgehoben wird – zum Beispiel durch expliziten Verweis auf die Fachlichkeit des Bereichs wie bei Klarheit und Strukturiertheit (ebd., S. 191) oder durch Rekurs auf domänenspezifisch ausformulierte Standards wie bei der Kompetenzorientierung (ebd., S. 240ff.) –, als beispielsweise im Bereich des *lernförderlichen Klimas*. Dies soll nachfolgend durch eine Darstellung der wesentlichen Aspekte der genannten Qualitätsbereiche verdeutlicht werden.

Sprachliche Äußerungen von Lehrkräften „müssen ... so *klar und verständlich* präsentiert und *strukturiert* werden, dass sie auf Schülerseite Lernprozesse anregen“ (ebd., S. 190). Die Bedeutung dieses Qualitätsmerkmals lässt sich auch aus empirischer Sicht bestätigen (ebd.). Zur begrifflichen Klärung bietet Helmke diese Unterscheidungen an:

- „Klarheit hat vier Komponenten: *akustisch* (Verstehbarkeit), *sprachlich* (Prägnanz), *inhaltlich* (Kohärenz) und *fachlich* (Korrektheit).
- Klarheit ist eher *sender-*, Verständlichkeit eher *empfängerbezogen*.
- Strukturiertheit hat zwei Bedeutungen: Aus *gedächtnispsychologischer* Sicht umfasst sie alle Merkmale des Informationsangebotes, die darauf abzielen, den Aufbau einer gut organisierten Wissensbasis zu unterstützen. Aus *didaktischer* Perspektive heißt Strukturiertheit, dass der Unterricht so geplant und sequenziert ist, dass er dieses Ziel erreicht.“ (Ebd., S. 191)

Zur Klarheit und Strukturiertheit gehört nach Helmke auch die *Verständlichkeit* (ebd., S. 196). Unter Rückbezug auf das *Hamburger Verständlichkeitskonzept* (Langer, Schulz von Thun & Tausch, 1974) werden hierunter Aspekte wie „*Einfachheit Kürze/Prägnanz Ordnung/Gliederung [und] [z]usätzliche Stimulanz [wie Alltagsbeispiele]*“ (Helmke, 2014, S. 196f.) gefasst. Unter dem Begriff der *Aktivierung* werden unterschiedliche Konzepte verstanden, wobei Helmke für die *kognitive Aktivierung* als „Aktivierung im Sinne der

Selbststeuerung des Lernens durch den Einsatz von Lernstrategien und Methoden eigenverantwortlichen Lernens“ (ebd., S. 205) unter Bezugnahme auf Hattie (2009) die aus empirischer Sicht hohe Wirksamkeit von Lernstrategien sowie das „Feedback zum aktuellen Lern- und Leistungsstand“ (Helmke, 2014, S. 214) betont. Beim Merkmal hoher *Motivierungsqualität* des Unterrichts wird das Potenzial kognitiver Konflikte als „Motor für das Lernen“ (ebd., S. 225) hervorgehoben, da „[d]as neu zu erwerbende Wissen ... zunächst in die vorhandene Wissensstruktur einzugliedern versucht [wird] (Assimilation)“ (ebd.) und anschließend, „[w]enn Elemente des neu zu erwerbenden Wissens mit dem vorhandenen Vorwissen unverträglich sind, ... Korrekturprozesse in Gang [kommen]: Die Wissensstrukturen selbst werden geändert (Akkommodation) [sic]“ (ebd.). Unter der Qualitätsdimension *Schülerorientierung* versteht Helmke den Anspruch, „sich in der Planung und Durchführung des Unterrichts konsequent am Schüler zu orientieren, an seinen Interessen, seinem Vorwissen, seiner Aktivierung etc.“ (ebd., S. 237). Die *Angebotsvariation* als unterrichtliches Qualitätsmerkmal wurde bereits erläutert (vgl. Kapitel 2.2.1), sodass abschließend zur *Kompetenzorientierung* auf die besondere Bedeutung der Orientierung an einem Lern-Outcome für die Planung von Unterricht hingewiesen werden kann, da Unterricht in diesem Sinne ‚von hinten‘ geplant werden muss (Helmke, 2014, S. 240ff.).

Eine andere Möglichkeit, sich qualitativen Merkmalen von Unterricht zu nähern, besteht in der Analyse von Standardformulierungen zum Lehrerberuf. So benennen die fächerübergreifenden *Kernstandards*²⁴ (*core propositions*) des *National Board für Professional Teaching Standards* (NBPTS, 2002) für guten Unterricht fünf Merkmale hoher Lehrkompetenz, die teilweise ähnliche Aspekte beinhalten wie die erläuterten Basisdimensionen:

- „1. Teachers are committed to students and their learning.
2. Teachers know the subjects they teach and how to teach those subjects to students.
3. Teachers are responsible for managing and monitoring students learning.
4. Teachers think systematically about their practice and learn from experience.
5. Teachers are members of learning communities.“ (S. 3f.)

Insbesondere in den ersten drei Punkten zeigen sich Ähnlichkeiten zu den bereits aufgeführten Übersichten über Qualitätsmerkmale. Es wird gefordert, dass Lehrer/-innen sich über die folgenden Aspekte bewusst sind: „preconceptions and background knowledge that students

²⁴ Es wird in Anlehnung an Helmke (2014) davon ausgegangen, dass „Standards ... nichts anderes als erwartete Kompetenzen [sind]“ (S. 142).

typically bring to each subject and ... strategies and instructional materials that can be of assistance“ (ebd., S. 3). Weiterhin wird Folgendes als ihre Aufgabe gesehen: „Accomplished teachers create, enrich, maintain and alter instructional settings to capture and sustain the interest of their students and to make the most effective use of time“ (ebd.). Die Dimensionen der individuellen Unterstützung der Lernenden sowie eine effektive Klassenführung lassen sich im Vergleich zu den oben genannten Basisdimensionen auch in diesen Beschreibungen erkennen.

Der stärker personenzentrierte Ansatz der bereits in Kapitel 2.2.1 vorgestellten Standardformulierungen zur Lehrerbildung der KMK (2004) wird in der Annahme deutlich, dass „[d]ie Qualität einer guten Schule und die Wirksamkeit eines guten Unterrichts ... entscheidend durch die professionellen und personalen Fähigkeiten von Lehrerinnen und Lehrern geprägt [werden]“ (S. 5); die Standards sollen hier nicht noch einmal aufgeführt werden. Im Hinblick auf die Fragen, inwieweit Aspekte hoher Unterrichtsqualität in den Kompetenzprofilen der Standardformulierungen enthalten sind, stellen A. Helmke, T. Helmke und Schrader (2007) fest:

- „Einige Punkte sind buchstäblich deckungsgleich, so z.B. die Überprüfung der Qualität des eigenen Lehrens und die Nutzung von Unterrichtsfeedback, das Angebot einer Vielfalt von Lernformen und die Motivierung.
- Bestimmte Aspekte der Unterrichtsqualität kommen nicht explizit vor, z.B. Punkte wie Klarheit/Strukturiertheit/Verständlichkeit/Passung/effiziente Klassenführung – sondern verstecken sich hinter der Generalklausel ‚gestalten Lehr-Lernprozesse unter Berücksichtigung der Erkenntnisse über den Erwerb von Wissen und Fähigkeiten‘.“ (S. 60f.)

Erwartungsgemäß lassen sich die zunächst überfachlichen Qualitätsmerkmale in dem fachspezifischen Kompetenzprofil der Standards für die Lehrerbildung für die sozialwissenschaftliche Fächerdomäne in noch geringerem Ausmaß identifizieren. Die Merkmale *Schülerorientierung*, *Angebotsvariation*, *Kompetenzorientierung* sowie *Umgang mit Heterogenität* werden in der Kompetenzerwartung allenfalls auf basalem Niveau angesprochen; demnach sollten Lehrer/-innen „exemplarisch fachliche Lehr- und Lernprozesse schüler- und problemorientiert diagnostizieren, analysieren, auch für heterogene Lerngruppen planen und arrangieren sowie Unterrichtsversuche im Fach evaluieren [können]“ (KMK, 2015, S. 48) sowie „über erste reflektierte Erfahrungen in der kompetenzorientierten Planung und Durchführung von Fachunterricht [verfügen] und ... Grundlagen der Leistungsdiagnose und -beurteilung im Fach [kennen]“ (ebd.). Eine weitere fachspezifische Rezeption von Qua-

litätsmerkmalen ist darüber hinaus nicht zu erkennen und es kann – wie es Helmke et al. (2007) für die allgemeinen Bildungsstandards formulieren – festgestellt werden, dass „eine starke Orientierung auf Wirkungen vor[herrscht]“ (S. 61).

Die von Oser (2001, S. 230ff.) formulierten 88 Standards zur Lehrerbildung fasst Helmke (2014) unter dem Gesichtspunkt ihrer Wichtigkeit für die Unterrichtsqualität in diesen Aspekten zusammen:

- „(A) *Lehrer-Schüler-Beziehungen*
 - (B) *Schülerunterstützendes Handeln und Diagnose*
 - (C) *Bewältigung von Disziplinproblemen und Schülerrisiken*
 - (D) *Gestaltung und Methoden des Unterrichts*
 - (E) *Zusammenarbeit in der Schule*
 - (F) *Allgemeine und fachdidaktische Standards*“
- (S. 151ff.; vgl. auch Oser, 2001, S. 230ff.)

Bis hierher lässt sich zusammenfassen, dass in unterschiedlichen Merkmalsbeschreibungen von Unterrichtsqualität mit dem Fokus auf den Aspekten, die die Qualität des *Unterrichtens* betreffen, teilweise Übereinstimmungen zu identifizieren sind. So wird die Orientierung an den Lernenden überall aufgeführt; ebenso sind unterrichtsmethodische Aspekte genannt. In einer noch allgemeinen und mehr das *Lehren* fokussierenden Formulierung scheint Unterricht also – unter der Prämisse des Ziels, Lernprozesse zu induzieren (vgl. Kapitel 2.1) – dann erfolgreich zu verlaufen, wenn es der Lehrkraft gelingt, adaptiv auf das individuelle (Fehl-)Verständnis der Schüler/-innen bezogene, kognitiv aktivierende sowie klare und strukturierte Impulse an diese zurückzumelden, die auf den Kompetenzaufbau abzielen. In einer Synopse theoretischer Annahmen sowie empirischer Befunde kommt Lindmeier (2011) für die Mathematik zu einem ähnlichen Schluss und formuliert zusammenfassend diese Aspekte erfolgreichen Unterrichtens: „1. correctness of content, 2. conceptions of students, 3. conceptual understanding, 4. connectedness, and 5. consistency and clarity“ (S. 101).

Inwieweit sich solche zunächst fachunspezifischen Merkmale der Unterrichtsqualität aus politikdidaktischer Sicht reflektieren lassen, soll im Folgenden dargestellt werden, da davon ausgegangen werden kann, dass sich diese auch in den Kompetenzfacetten des Wissens über Instruktions- und Vermittlungsstrategien sowie des Wissens über (Fehl-)Vorstellungen der Lernenden als qualitative Merkmale identifizieren lassen können.

Das Qualitätsmerkmal *kognitive Aktivierung* ist bislang in der Politikdidaktik erst vereinzelt explizit rezipiert worden. In einer pilotierenden *Smallscale*-Studie zum Konzeptaufbau untergliedern Manzel und Gronostay (2013) das Merkmal der kognitiven Aktivierung in diese vier Unterfacetten:

- „a) Authentizität (situiertes und exemplarisches Lernen, Fallprinzip)
- b) Fachkonzepte: ‚Politik kommt in Politik-Stunden vor‘
- c) Eigenaktivität, kooperatives Lernen und instruktionale Unterstützung
- d) Diskursivität (Redeanteile, Elaborationsdichte)“ (S. 200f.)

Die deskriptiven Ergebnisse „zeigen ... [in den] videografierten Politikstunden vielfach nur eine geringe Vernetzung der fachlichen Begriffe und mit einem Prozent nur äußerst selten ein komplexes Vernetzungsniveau beim Aufbau eines Fachkonzepts bei den Lernenden“ (ebd., S. 213). Die Ergebnisse einer Studie von Weißeno und Landwehr (2015) zur Perzeption von Politikunterricht zum Themenfeld *Europäische Union* weisen in den erhobenen Unterrichtsstunden nach, dass, „[o]bgleich der Politikunterricht schwach strukturiert ist, ... sich die wahrgenommene Aktivierung positiv auf die Leistung aus[wirkt]“ (S. 106). Die zentrale Stellung der Dimension kognitiver Aktivierung für die Leistung der Schüler/-innen wird hervorgehoben (ebd.; vgl. auch Detjen et al., 2012, S. 115). Auf mögliche Ansatzpunkte für die politikdidaktische Theoriebildung weist auch D. Neumann (2014, S. 55) hin, sodass in Verbindung mit den vorliegenden empirischen Befunden der fachspezifischen Unterrichtsforschung die Wichtigkeit des Aspekts der kognitiven Aktivierung als Merkmal hoher Fachunterrichtsqualität angenommen werden kann.

Die *Schülerorientierung* erhält in den oben aufgeführten Merkmalsbeschreibungen zwar unterschiedliche Ausformulierungen, jedoch wird im Wesentlichen die Auffassung geteilt, „dass Schüler, unabhängig von Lernen und Leistung, als *Person* ernst genommen und wertgeschätzt werden“ (Helmke, 2014, S. 236), was „viel mit dem *affektiven* Aspekt der Lehrer-Schüler-Beziehung und dem *Wohlbefinden* von Schülerinnen und Schülern zu tun [hat]“ (ebd.). Dieses Merkmal von Unterrichtsqualität zeigt dabei eine inhaltliche Nähe zum didaktischen Prinzip der Schüler- oder auch Adressatenorientierung²⁵. Sander (2008) fasst zusammen, dass Schülerorientierung für die Planung und Gestaltung des Politikunterrichts impliziert, dass „Lerngegenstände ... so ausgewählt und strukturiert werden [sollten], dass sie

²⁵ Der Begriff der Schülerorientierung ist im Bereich schulischer politischer Bildung gebräuchlicher (Sander, 2008, S. 191).

an Vorwissen und Voreinstellungen der Adressaten über Politik anschlussfähig sind und dass die Adressaten als Subjekte ernst genommen sowie ihre Lebenserfahrungen und Lerninteressen berücksichtigt werden [Umformatierung v. Verf.]“ (S. 191). Übereinstimmend sagt auch Detjen (2007), dass „[d]er didaktische Grundsatz betont, dass die Lernenden als Subjekte der politischen Bildung ernstgenommen werden müssen“ (S. 330). Reinhardt (2012b, S. 202f.) führt die Schülerorientierung zwar nicht als politikdidaktisches Prinzip an, hebt jedoch ebenfalls dessen Wichtigkeit für die Planung unterrichtlicher Lehr-Lern-Prozesse hervor. Bei Schmiederer (1977, S. 78ff.) fußt eine didaktische Gesamtkonzeption auf dem Aspekt der Orientierung am Interesse der Schüler/-innen, wobei Schelle (2005) darauf hinweist, dass sich die Rezeption dieses Prinzip über die Jahre hinweg geändert haben dürfte:

„Es geht heute nicht mehr bloß darum, dass die Lehrerinnen und Lehrer als die Adressaten der universitär organisierten Ausbildung in der Lage sein müssen, die Schülerinnen und Schüler als Adressaten der schulischen politischen Bildung zu verstehen, deren Verstehensbedürfnisse und Interessen aufzunehmen und sie zum politischen Lernen zu animieren, sondern auch darum, dass die Schülerinnen und Schüler selber befähigt werden müssen, ein reflexives Verhältnis zum Verstehen von Selbst und Welt zu entwickeln.“ (S. 87)

Im Gegensatz zur Betonung der Anschlussfähigkeit an das Vorwissen der Lernenden bei Sander (2008) weist Detjen (2007) darauf hin, dass „[d]ie Forderung, dass Lernprozesse anschlussfähig sein müssen an Vorwissen und Voreinstellungen der Lernenden, [zwar] ... mit der Schülerorientierung [harmoniert], ... selbst aber kein Ausfluss des politikdidaktischen Prinzips [ist]“ (S. 330).

Es zeichnen sich zwei Dimensionen der Schülerorientierung in der politischen Bildung ab: eine affektiv-emotionale sowie eine, die eine Anpassung der Lernangebote an die (Fehl-)Vorstellungen der Schüler/-innen betont. Erstere zeigt sich inhaltlich anschlussfähig an die bisher aufgeführten Auffassungen von Schülerorientierung (Helmke, 2014, S. 236ff.). Die affektiv-emotionale Dimension dürfte im Politikunterricht in der direkten Kommunikation zwischen Lehrenden und Lernenden sowie auch innerhalb der Klasse wichtig sein, da „sich Politik ... von anderen Fächern auch dadurch [unterscheidet], daß [*sic*] in diesem Unterricht der Kommunikation eine besondere Bedeutung zukommt“ (Ackermann, 2010, S. 133) und „im Kommunikationsprozeß [*sic*] nicht nur Wissen, sondern auch Sozialerfahrung und soziale Kompetenz vermittelt [wird]“ (ebd.). Darüber hinaus ist eine adäquate Unterrichtskommunikation auch ein wichtiges Korrektiv bei der Umsetzung des Beutelsbacher Konsenses im Unterrichtsverlauf und hier insbesondere in der Anforderung an

Lehrende des Fachs Politik, dass sie die eigene Urteilsbildung der Schüler/-innen nicht behindern dürfen. Die Wichtigkeit einer adäquaten Unterrichtskommunikation wird übereinstimmend von weiteren Autoren hervorgehoben (Massing, 2005; Petrik, 2014).

Eine Orientierung an den Lernenden als Unterrichtsqualitätsmerkmal nach der zweiten Auffassung hat besondere Bedeutung für den Kompetenzerwerb, speziell für den Aufbau konzeptuellen Politikwissens (Detjen et al., 2012, S. 29ff.). Hierbei „impliziert Kompetenzorientierung *Schülerorientierung* in dem Sinne, dass auf zentrale, langfristig aufgebaute Lernergebnisse fokussiert wird, also auf das, was die Schülerinnen und Schüler nach der Lernphase tatsächlich beherrschen“ (Hofer, 2012, S. 18). Mit Bezug auf Theorien zum *conceptual change* betonen Weißeno et al. (2010) zur Entwicklung der Kompetenzdimension *Fachwissen*, dass Wissen ... im Unterricht gut strukturiert zu präsentieren [ist], und [dass] das Wissen ... von den Lernenden mit ihrem Vorwissen vernetzt werden können [muss]“ (S. 21). Der Orientierung am Fehlverständnis der Schüler/-innen kommt dabei nicht nur in der Planung von Lernangeboten eine besondere Bedeutung zu (Manzel & Neumann, 2015, S. 72f.), sondern auch in den Rückmeldungen der Lehrkräfte an die Schüler/-innen (für die Domäne der Naturwissenschaften vgl. Smith & Neale, 1989, S. 3f.). Lindmeier (2011) fasst diesen Aspekt erfolgreichen Unterrichtens so zusammen: „Students‘ conceptions, including errors and misconceptions, are the starting point for instruction“ (S. 101). Schülerorientierung scheint somit in beiden ausgeführten Dimensionen ein wichtiges Merkmal guten Politikunterrichts zu sein, obgleich eine systematisch-empirische Auseinandersetzung mit diesem Aspekt bislang noch aussteht.

Für eine an Bildungsstandards und damit an dem Outcome des Unterrichts orientierte Planung von Lernangeboten wird *Kompetenzorientierung* als Merkmal hoher Fachunterrichtsqualität betont (Detjen et al., 2012, S. 111ff.; Weißeno et al., 2010, S. 22f.; für eine Kritik vgl. Sander, 2011, 2013). Als eine zentrale Dimension domänenspezifischer Kompetenzen von Lernenden wird dabei das bereits angesprochene Fachwissen in den Dimensionen des Fakten- und Konzeptwissens angesehen, weil Schüler/-innen, „[w]enn ... [sie] über konzeptuelles Politikwissen verfügen, ... in späteren Anwendungssituationen, sei es als Schüler oder als Bürger, zu kontextadäquaten Lösungen kommen [können], obwohl sie eine konkrete Erfahrung mit der erforderlichen Lösung noch nicht gemacht haben“ (Detjen et al., 2012, S. 29). Erfolgreicher Politikunterricht sollte sich im Hinblick auf das

Qualitätsmerkmal der Kompetenzorientierung insbesondere auch dadurch auszeichnen, dass es der Lehrperson gelingt, die Schüler/-innen nicht nur beim Erwerb singulären Faktenwissens, sondern beim Aufbau konzeptuellen Wissens zu unterstützen. Dies sollte sich auch in Rückmeldungen der Lehrkraft zeigen (Lindmeier, 2011, S. 101).

Dass diese dem Anspruch fachlicher Korrektheit genügen müssen, erscheint zunächst redundant, ist aber – auch vor dem Hintergrund des hohen Anteils fachfremd erteilten Politik-/SoWi-Unterrichts (Landesinitiative FÜR politische Bildung, 2006) – durchaus hervorzuheben. Im Zusammenhang mit der Darstellung der eng miteinander verbundenen Qualitätsmerkmale *Klarheit und Strukturiertheit* betont Helmke (2014) daher auch, dass „[f]achliche Korrektheit ... zwar nicht alles [ist], aber ohne fachliche Richtigkeit und Genauigkeit ... alles andere nichts [ist]“ (S. 199). Es kann davon ausgegangen werden, dass Rückmeldungen von Politiklehrkräften inhaltlich verständlich und dabei stets fachlich korrekt sein müssen, um bei den Lernenden nicht dem Aufbau von Fehlvorstellungen Vorschub zu leisten. Gleichzeitig ist zu betonen, dass Zusammenhänge zwischen der Fachsprache und dem konzeptuellen Verständnis betont werden (für die mathematisch-naturwissenschaftliche Domäne vgl. Yore, Pimm & Tuan, 2007) und Lehrende darüber hinaus auch als Sprachmodell für Schüler/-innen fungieren (Helmke, 2014, S. 194f.), weshalb die Fachsprache als domänenspezifischer Aspekt sprachsensiblen Politikunterrichts zunehmend diskutiert wird (Oleschko, 2013). Es ist hierbei zu berücksichtigen, dass viele politische Fachbegriffe zudem alltagssprachlich anders konnotiert sind – zum Beispiel *Gewalt* im politikwissenschaftlichen Verständnis als *Staatsfunktion* versus Gewalt im alltagssprachlichen Sinne als *Gewalttätigkeit* (Weißeno et al., 2010, S. 70) – und dass ein sprach-sensibler Politikunterricht daher hohe Anforderungen an die Rückmeldungen der Lehrkräfte stellen dürfte. Schließlich ist die fachliche Korrektheit als Aspekt von Klarheit und Strukturiertheit auch deshalb hervorzuheben, weil die angesprochene Anpassung der Rückmeldung an die Verständnisschwierigkeit der Lernenden überhaupt nur dann gelingt, wenn die Lehrkraft den Fehler eindeutig und fachlich korrekt identifizieren kann. Zusammenfassend für dieses Merkmal von Unterrichtsqualität lässt sich aus politikdidaktischer Perspektive seine fachspezifische Bedeutung insbesondere im Aspekt der fachlichen Korrektheit unterstreichen. Allerdings sind hierzu im Bereich empirischer Fachunterrichtsforschung in der Politikdidaktik erhebliche Forschungslücken zu erkennen.

Die Bedeutung der *Motivierungsqualität* von Unterricht wird in der Politikdidaktik im Zusammenhang mit den erhofften Wirkungsweisen von Unterrichtsmethoden und fachdidaktischen Prinzipien thematisiert (Detjen, 2007, S. 320). Insbesondere das motivierende Potenzial kognitiver Konflikte (Helmke, 2014, S. 225) scheint für die Politikdidaktik im Rahmen fachdidaktischer Zugänge zur Unterrichtsgestaltung von Interesse zu sein, augenfällig beispielsweise im politikdidaktischen Prinzip der Konfliktorientierung (Reinhardt, 2012b, S. 76ff.).²⁶ Aber auch bei den fachdidaktischen Prinzipien der Kontroversität und der Problemorientierung wird davon ausgegangen, dass die aufgeworfenen Diskussionen und Problemlagen kognitive Konflikte im Sinne kognitiver Dissonanzen (Festinger, 2012) bei den Lernenden auslösen können (D. Neumann, 2014, S. 55), um sie somit zur aktiven Auseinandersetzung mit dem Stoff zu motivieren. Wenn sie zur ‚Auflösung‘ von Fehlkonzepten eingesetzt werden, lassen sich in Anlehnung an Theorien zum Conceptual-Change-Unterrichten derartige fachdidaktische Zugänge zum Unterricht auch als „Dissonance Strategies“ (Clement, 2008, S. 421) beschreiben. Umsetzungsmöglichkeiten solcher Strategien ergeben sich zum Beispiel durch „forms of ‚participative dissonance,‘ [sic] where information is provided that allows the student to discover a conflict with his or her own current model“ (ebd.). Dabei sind „[d]iscrepant events ... empirical experiments, data summaries, or demonstrations that provide data that could promote dissonance with students‘ preconceptions“ (ebd.). Es ist jedoch festzuhalten, dass eine systematische empirische Auseinandersetzung mit den Wirkungen von politikdidaktischen Prinzipien und Methoden als Strategien des Conceptual-Change-Unterrichtens bislang kaum vorhanden ist.

Die politikdidaktische Reflexion der hier aufgeführten Merkmale von Unterrichtsqualität zeigt, dass eine fachspezifische Rezeption teilweise bereits explizit erfolgt (kognitive Aktivierung, Kompetenzorientierung), zum Teil jedoch erst implizit – also nicht in Form einer Auseinandersetzung zum Zwecke der Bestimmung empirisch diagnostizierbarer Qualitätskriterien – stattfindet (Schülerorientierung, Klarheit und Strukturiertheit). Da diese Merkmale von vielen Autoren auch für den Politikunterricht als wichtig erachtet werden, wird im Folgenden davon ausgegangen, dass diese als „[p]erformance criteria“ (Lindmeier, 2011, S.

²⁶ Hierbei ist zu berücksichtigen, dass – zum Beispiel im unterrichtsmethodischen Zugang in Form einer Konfliktanalyse – im Sinne hoher Authentizität von gesellschaftspolitisch *realen* Konflikten ausgegangen wird (Reinhardt, 2012b, S. 81ff.); ob sich diese Konflikte auch als kognitive Konflikte ‚in den Köpfen‘ der Lernenden abzeichnen, ist damit noch nicht gesagt.

102) für die qualitative Diagnostik der politikdidaktischen Kompetenz als Referenzrahmen dienen können.

2.3.4 Politikdidaktische Kompetenzen im fachdidaktischen Diskurs

Eine dezidiert an der Modellierung von Lehrerkompetenzen orientierte Auseinandersetzung mit den Kognitionen von Politiklehrkräften findet erst seit einigen Jahren in der Politikdidaktik statt. Entsprechend disparat ist das Bild von strukturtheoretisch²⁷ und kompetenztheoretisch fundierten Konzeptualisierungen (Weschenfelder, 2014a, S. 52). Im Folgenden werden ausgewählte Überlegungen unterschiedlicher Autoren zum professionellen Wissen und Können von Politiklehrkräften dargestellt und dabei diejenigen Ansätze fokussiert, die sich auf einen Kompetenzbegriff beziehen. Gleichzeitig soll kritisch reflektiert werden, inwieweit sich die Ansätze an den zuvor dargestellten kompetenztheoretischen Grundlagen, die für eine an die Befunde anderer Domänen anschlussfähige und theoretisch begründete Konzeptualisierung professioneller Kompetenzen von Lehrkräften heranzuziehen wären, orientieren. Vor dem Hintergrund der leitenden Fragestellungen soll in diesem Kapitel daher das Verhältnis von Wissen und Können in den vorliegenden Ansätzen analysiert werden, sofern in diesen thematisiert.

Reinhardt (2012b) beschreibt als Anforderung an den Politiklehrer, dass dieser „Routinen des Lehrens und auch seine Fachwissenschaften (also die Bezugsdisziplinen für seine Unterrichtsfächer) in einem Ausmaß [beherrscht], dass sein beruflicher Alltag erfolgreich verläuft“ (S. 11), womit insbesondere beim Auftreten unterrichtsbezogener Probleme eine kritische Reflexionsfähigkeit über das eigene Handeln verbunden sei (ebd.), sowie an anderer Stelle, dass „[d]er professionelle Lehrer ... theoretische fachdidaktische Konstrukte für die Konstruktion seines Unterrichts [benutzt]“ (2009, S. 25), wodurch „[d]iese mögliche Theoretisierung ... den Professionellen vom Profi [unterscheidet], der sehr wohl erfolgreich Routinen reproduziert, dies aber auf intuitiver oder eingeübter Grundlage tut“ (ebd.). Ausgangslage für die Überlegungen zum Wissen von Politiklehrkräften ist eine Analyse zur

²⁷ Für eine Übersicht über strukturtheoretische Grundlagen der Professionalität von Lehrkräften, die im Folgenden unberücksichtigt bleiben muss, vgl. Helsper (2014); eine Zusammenschau aus politikdidaktischer Perspektive bietet Weschenfelder (2014a, S. 18ff.). Zur spezifischen Problematisierung von domänenspezifischen unterrichtlichen Anforderungen an das Handeln der Lehrpersonen aus Sicht einer strukturtheoretischen Professionstheorie vgl. May (2014).

Konfliktstruktur der widersprüchlichen Rollenanforderungen in Interaktionen zwischen Lehrenden und Lernenden. Das in Anlehnung an Baumert und Kunter (2006) formulierte Professionswissen soll bei Reinhardt (2009) inhaltlich durch fachdidaktische Prinzipien gefüllt werden, die „die Brücke zwischen Theorie und Praxis [bauen], ... in kurzer Form Ziele, Inhalte, Bedingungen und Methodik des Unterrichts [integrieren] und ... vier unterschiedliche Wissensformen in Beziehung [setzen]“ (S. 25). Diese vier Wissensformen werden dabei so ausformuliert:

„1. normatives Wissen: Werte und Überzeugungen vom Sinn und Ziel des Lernens (Demokratie-Lernen mit dem Ziel der Mündigkeit), 2. Fach(wissenschaft)liches Wissen: Die Gegenstände des Faches werden u. a. von den Sozialwissenschaften beschrieben und erklärt, 3. Alltagswissen: Alltägliche Zugänge der Lernenden sind andere als die der Fachwissenschaften oder der Experten, 4. Berufswissen: Lehrer verfügen über eine Fülle von Verfahren und Strategien, wie Lernprozesse organisiert werden können (von Methoden als Struktur einer ganzen Unterrichtsreihe bis zu Medien als Arbeitsmittel – vgl. Reinhardt / Richter (Hg.) 2007).“ (Ebd.)

Die Bedeutung des professionellen Wissens wird in diesem Entwurf hervorgehoben, da „[z]war ... die vierte Wissensform des Berufswissens dem Handeln am nächsten [ist], aber auch dieses Berufswissen ... theoretisches Wissen über Handeln [ist]“ (Reinhardt, 2012a, S. 41), weshalb „[r]eflexive Kompetenz ... unerlässlich [ist], damit professionell gehandelt werden kann“ (Reinhardt, 2009, S. 26). Fachdidaktische Prinzipien sollen Lehrkräften als Unterrichtsstrategien dienen und dadurch bei der Handhabung der skizzierten Konfliktlagen unterstützen. Die Auffassung, dass „[p]rofessionelles Lehrerhandeln ... Theorie und Praxis [verknüpft], indem fachdidaktisches Wissen und praktisches Können reflexiv vermittelt werden“ (Reinhardt, 2012a, S. 36), lässt vermuten, dass von einer gewissen Handlungswirksamkeit des politikdidaktischen Wissens über fachdidaktische Prinzipien ausgegangen wird. Es wird jedoch auch betont, dass dieses Wissen weiterhin reflexives Wissen im Sinne eines Wissens über Handeln sei (Reinhardt, 2009, S. 26). Dabei „ist aber nicht gesagt, dass das richtige Wissen, wie zu handeln sei, auch das Können verbürgt“ (ebd.), was – neben der Frage, was ‚richtiges Wissen‘ ist – offenlässt, aus welchen kognitiven Ressourcen sich das Unterrichtshandeln speist.

Dem Entwurf von Reinhardt (2009, 2012a) liegt kein Modell im Sinne der Darstellung in Kapitel 2.3.2 zugrunde, weshalb auch „[e]ine Differenzierung verschiedener Kompetenzdimensionen (beispielsweise Wissen und Überzeugungen) und die Aufnahme weiterer wesent-

licher Kompetenzen (z. B. motivationale Orientierungen) fehlen“ (Weschenfelder, 2014a, S. 53f.). Dies erschwert eine systematische Zuordnung dieser Wissensdimensionen zu einem übergeordneten Kompetenzstrukturmodell sowie eine Operationalisierung der Konstrukte und deren Erhebung. Unklar bleibt auch die Strukturierung der Dimensionen der vier vorgeschlagenen Wissensformen. Ob beispielsweise unter der *„Dimension des normativen Wissens“* (Reinhardt, 2012a, S. 35) im engeren Sinne nicht Wissen, sondern andere Kompetenzaspekte wie professionelle Überzeugungen subsumiert werden, ist nicht eindeutig geklärt. Weiterhin wird im Entwurf von Reinhardt (2009, S. 24) zwar auf andere Autoren wie Baumert und Kunter (2006) rekurriert, jedoch werden zentrale Begriffe wie der des Professionswissens in einem divergenten Verständnis verwendet. So wären nach der Modellierung von Baumert und Kunter (2006) sowie der damit verbundenen Wissenstaxonomie von Shulman (1986, 1987) fachdidaktische Prinzipien allenfalls ein Aspekt einer Facette des fachdidaktischen Wissens von Lehrkräften (vgl. auch Kapitel 2.2.1), während solche Prinzipien bei Reinhardt (2009, S. 25) mit dem Professionswissen begriffsgleich erscheinen.

Sander (2008) formuliert unter der Prämisse, dass „[i]n der politischen Bildung ... Lehrende Menschen dabei [unterstützen], sich ihr eigenes Bild von Politik zu erarbeiten, zu erweitern und ggf. auch zu verändern, also ein gutes Stück weit ihre politische Sozialisation selbst in die Hand zu nehmen“ (S. 169), diese Anforderungen an Politiklehrkräfte: Schaffung von Lernumgebungen²⁸, Diagnostik, Interventionen zur Anregung von Lehr-/Lern-Prozessen, Beurteilen/Bewerten sowie Schul-/Organisationsentwicklung (ebd.). Einen weiteren möglichen Zugang zu den Problemstellen des Lehrens im Sinne der Rekonstruktion unterrichtlicher Anforderungssituationen sieht Zurstrassen (2009) in den „von Sander herausgearbeiteten ‚Fallen‘, die beschreiben, woran politisches Lernen scheitert“ (S. 40). Im weiteren Verlauf wird bei Sander (2008, S. 169ff.) auf die von Girmes (2006) vorgeschlagene Systematisierung von zehn „Qualifizierungsaufgaben ...[,] mit denen sich professionell Lehrende zumindest erfolgreich auseinandergesetzt haben sollten“ (S. 27)

²⁸ Der Begriff der Lernumgebung wird in diesem Zusammenhang in einem konstruktivistisch-didaktischen Verständnis verwendet, um zu betonen, dass Lernen ein aktiver, selbst gesteuerter Prozess ist, den Lehrende nicht ‚von außen‘ lenken können, sondern allenfalls durch die Schaffung passender Lernangebote (d. h. einer kognitiv anregenden Lernumgebung) versuchen können, positiv zu beeinflussen (zur Perspektive des Konstruktivismus in der Politikdidaktik vgl. Sander, 2008, S. 162ff.; Scherb, 2002). Nach einem erweiterten Verständnis gehören damit auch die Lehrpersonen selbst zur Lernumgebung (Sander, 2008, S. 173).

rekurriert; die Kompetenzen, die Lehrer/-innen zur Bewältigung der unterschiedlichen Aufgaben ausbilden sollen, umfassen:

*„Diagnostische Kompetenz
Institutionelle Kompetenz
Curriculare Kompetenz
Methodische Kompetenz
Personell-kulturelle Kompetenz
Reflexive und evaluative Kompetenz“ (Ebd.)*

Es kann angemerkt werden, dass zum einen die von Girmes (ebd., S. 25ff.) vorgeschlagenen Kompetenzen wenig anschlussfähig an einen Kompetenzbegriff sind, wie er in Kapitel 2.3.1 erläutert wurde, da wesentliche Kompetenzdimensionen wie motivationale Orientierungen unberücksichtigt bleiben. Zum anderen bleibt unklar, in welchem strukturellen Verhältnis der Vorschlag von Girmes (ebd.) zu den von Sander (2008) formulierten Anforderungsdimensionen steht, da auch eine domänenspezifische Rezeption ausbleibt.

Bei Langner (2013) erfolgt die Beschreibung der Fähigkeiten und Fertigkeiten für ein professionelles Handeln von Politiklehrkräften entlang dreier Merkmalsdimensionen, die in „die berufsethische, die fachlich standardisierte sowie die funktionsbezogene Dimension“ (S. 135) aufgegliedert werden. In der ersten Dimension sind im Wesentlichen Aspekte professioneller Überzeugungen und Werthaltungen angesprochen (ebd.), während in der zweiten das eigentliche Professionswissen von Lehrkräften enthalten ist (ebd., S. 136f.). Abgeleitet aus dem normativen Ziel, „Lernende zu der Beschäftigung mit sozialwissenschaftlichen Fragestellungen aus einer politischen Perspektive heraus anzuleiten und anzuregen“ (ebd., S. 137), werden unter dem professionellen Wissen unter anderem Kenntnisse zu fachdidaktischen Prinzipien, den auszubildenden fachspezifischen Kompetenzen von Lernenden sowie deren Einstellungen subsumiert (ebd., S. 136ff.). Das Vorhandensein adäquaten Professionswissens wird als notwendig erachtet, um professionell handeln zu können (dritte Merkmalsdimension), worunter zur Bewältigung von vier – in Anlehnung an die Standardformulierungen der KMK (2014) formulierten – sogenannten Tätigkeitsprofilen folgende Fähigkeiten verstanden werden: „1. Politische Bildungsarrangements inszenieren und politisches Lernen anregen ... 2. Politische Lernzuwächse erfassen sowie fördern und Feedback geben ... 3. Mit Heterogenität umgehen und zum politischen Urteilen wie auch Handeln befähigen ... 4. Politische Bildungstätigkeit

weiterentwickeln und Bildungssettings innovieren ... [Umformatierung v. Verf.]“ (Langner, 2013, S. 140).

Anzumerken zum Entwurf von Langner (ebd.) ist, dass er neben dem Rückbezug auf die Standardformulierungen der KMK (2014) keinen weiteren theoretischen Bezugsrahmen der Kompetenzmodellierung darlegt und insofern nicht deutlich wird, in welchem strukturellen Gesamtzusammenhang die vorgeschlagenen Aspekte professionellen Handelns von Politiklehrkräften wie Wissen, Können, motivationale Orientierungen und Überzeugungen zueinander stehen. So wird hinsichtlich des Verhältnisses von Wissen und Können „die Dichotomie von politikdidaktischem Denken und politikdidaktischem Handeln“ (Langner, 2013, S. 141) zwar abgelehnt und es wird von einem Wechselwirkungsverhältnis zwischen Theorie und Praxis ausgegangen (ebd.); es bleibt jedoch offen, inwiefern theoretisches Wissen handlungsrelevant wird, da in der zweiten Dimension (Professionswissen) wesentliche Merkmale des Tätigkeitsprofils von Politiklehrkräften (dritte Dimension professionellen Handelns) – wie zum Beispiel die Fähigkeit, „sich an der Planung und Umsetzung von (neuen) politischen Bildungssettings [zu] beteiligen“ (ebd., S. 140) – nicht enthalten sind.

Der Entwurf von Petrik (2009, 2012) erfolgt anhand politikdidaktischer Problemstellen, die die Rekonstruktion der Handlungsanforderungen an Politiklehrkräfte ermöglichen sollen. Dabei umfasst „[d]as *Gegenstandsproblem* ... die Schwierigkeit von Lehrern und Lehrerinnen, das Politische als abgrenzbares, übersichtliches und überzeitliches Phänomen darzustellen“ (Petrik, 2012, S. 71); „[d]as *Brückenproblem* ... die methodische Schwierigkeit von Lehrern und Lehrerinnen, eine nachhaltige Begegnung zwischen Lernenden und Gegenstand anzubahnen, die an Gesellschaftsbilder, Alltagstheorien und Erfahrungen von Lernen anknüpft“ (ebd., S. 72), und „[d]as *Aushandlungsproblem* ... die interaktiven Schwierigkeiten zwischen LehrerInnen und SchülerInnen sowie der SchülerInnen untereinander, die häufig durch einseitige Medienwahl und Medien-Deutung verstärkt werden“ (ebd.). Diese Problemstellen lassen sich nach Petrik (2009, 2012) im didaktischen Dreieck verorten, das insofern nicht nur die Implikationszusammenhänge zwischen den drei Eckpunkten Lehrer/-in, Schüler/-in und Lerngegenstand, sondern auch zwischen den politikdidaktischen Problemstellen verdeutlichen soll. Die Begründung der Aspekte politikdidaktischer Kompetenzen, die „den professionellen Umgang mit diesen

unhintergehbaren Problemstellen [definieren]“ (Petrik, 2012, S. 72), erfolgt in Anlehnung an Reinhardts (2009) Entwurf in Verbindung mit der Formulierung fachdidaktischer Kompetenzbereiche der GFD (2005). Petrik (2009) schlägt diese vier politikdidaktischen Kompetenzaspekte vor:

- „1. *Didaktische Analysekompetenz* (theoriegeleitete fachdidaktische Reflexion)
2. *Didaktische Inszenierungskompetenz* (fachbezogenes Unterrichten)
3. *Didaktische Aushandlungskompetenz* (fachbezogene(s) Kommunikation & Unterrichten)
4. *Didaktische Diagnosekompetenz* (fachbezogenes Diagnostizieren, Beurteilen, Entwickeln und Evaluieren)“ (S. 68)

Bezogen auf die ersten beiden Aspekte werden Fähigkeiten des souveränen Umgangs mit politikdidaktischen Konzeptionen, die „fachdidaktische Prinzipien aus[bilden]“ (Petrik, 2012, S. 74) beschrieben. Es zeigt sich eine Verbindung zum Systematisierungsansatz Reinhardts (2009, 2012a), wobei Modifikationen – beispielsweise hinsichtlich der Zuordnung von Methoden zu den Prinzipien – vorgenommen werden (Petrik, 2012, S. 75). Schließlich entwirft Petrik (ebd., S. 77ff.) unterschiedliche Unterrichtsleitbilder, um hieran mögliche Konflikte zwischen den bzw. innerhalb der Überzeugungen von Lehrkräften zu verdeutlichen. Nach dieser Auffassung wird unterrichtliches „[d]idaktisches Handeln ... primär aus zwei Quellen gespeist: Didaktisches Theoriewissen vermischt sich mit beständigen didaktischen Alltagsvorstellungen, die als wertgebundener Wahrnehmungsfilter fungieren und institutionellen Handlungsbedingungen, so dass ein direkter Wissenstransfer unmöglich ist“ (Petrik, 2009, S. 63). Handlungsnahes Wissen, das „sich selbstverständlich erst in der Praxis voll entwickeln kann“ (ebd., S. 76), wird insbesondere unter dem Aspekt der didaktischen Aushandlungskompetenz zusammengefasst.

Die politikdidaktische Kompetenz kann bei Petrik (2009, 2012) als Amalgam aus Wissen und Können angesehen werden; unklar bleibt jedoch, welche Rolle motivationale und volitionale Aspekte beim unterrichtlichen Handeln spielen, da – wohl auch wegen der dezidiert kritischen Distanz zum Modell professioneller Handlungskompetenzen bei COACTIV (Petrik, 2009, S. 67) – keine Verortung in einem übergeordneten Strukturmodell expliziert wird (vgl. auch Weschenfelder, 2014a, S. 56). Das ‚reflexive Moment‘, in dem „[e]inmal ein inhaltliches Wissen um typische Antinomien des Lehrerhandelns und zum anderen ein methodisches Interpretationswissen“ (Petrik, 2009, S. 63) enthalten ist, kann

dementsprechend auch nicht vollumfänglich erklären, warum beispielsweise theoretisches fachdidaktisches Wissen in unterrichtlichen Anforderungssituationen träge bleiben kann.

An dieser Stelle kann aus kompetenztheoretischer Perspektive zusammengefasst werden, dass die bislang referierten Entwürfe von Reinhardt (2009, 2012a) und Petrik (2009, 2012) zwar auf vorhandene generische Kompetenzmodellierungen (zum Beispiel des COACTIV-Forschungsprogramms) rekurren, dies jedoch teilweise in kritischer Distanz oder lediglich auf basalem Niveau. So lässt die Verwendung des Begriffs des Professionswissens bei Reinhardt (2009, S. 24) eine Nähe zur Konzeptualisierung von Baumert und Kunter (2006, 2011a) erkennen, jedoch werden unter dieser Kompetenzdimension auch weitere Kompetenzaspekte wie Motivation und Überzeugungen zusammengefasst, was insofern nicht anschlussfähig an die kompetenztheoretischen und kognitionspsychologischen Grundlagen vorliegender Modellierungen erscheint. Die zum Teil unklare wissenschaftstheoretische Verortung der Entwürfe von Ansätzen von Langner (2013), Petrik (2009, 2012), Reinhardt (2009, 2012a) und Sander (2008) lässt sich auch daran festmachen, dass die genannten Autoren nicht explizieren, welches Kompetenzkonstrukt ihren Konzeptualisierungen zugrunde liegt.

Politikdidaktische Prinzipien werden bei Petrik (2009, 2012) und Reinhardt (2009, 2012a) als wesentlicher Aspekt des professionellen Wissens von Politiklehrkräften angesehen, wobei Reinhardt (2009, S. 25) diesen Prinzipien aufgrund der angesprochenen Subsumtion unterschiedlicher Kompetenzaspekte unter einer zentralen Kompetenzdimension *Professionswissen* einen noch umfassenderen Charakter zuschreibt als Petrik (2009, S. 70f.), bei dem die fachdidaktischen Prinzipien als ‚Werkzeuge‘ für Lehrkräfte eher zwei von vier Kompetenzaspekten zugeordnet werden können. Insbesondere bei Reinhardt (2009, 2012a) erscheint das Verständnis zum Professionswissen wenig anschlussfähig zum Beispiel an die Taxonomie von Shulman (1986, 1987), wie sie in der deutschsprachigen Forschung rezipiert wird (vgl. auch Kapitel 2.2). Schließlich lässt sich hinsichtlich der Frage nach den Ansätzen zur Operationalisierung der konstruierten Kompetenzdimensionen und -facetten in den Entwürfen von Langner (2013), Petrik (2009, 2012), Reinhardt (2009, 2012a) und Sander (2008) problematisieren, dass nicht ausformuliert wird, wie sich die vorgeschlagenen Wissens- und Kompetenzbereiche in Test-Items überführen lassen können, was jedoch für handlungsnahe Kompetenzbereiche wie die *didaktische Aushandlungskompetenz* bei Petrik (2012, S. 72)

notwendig wäre, da sich diese wahrscheinlich nicht im Rahmen eines traditionellen schriftlichen Tests erheben ließe.

Das in der PKP-Studie des Forschungsverbundes Karlsruhe-Göttingen von Oberle, Weschenfelder und Weißeno (2012) entwickelte domänenspezifische Strukturmodell professioneller Kompetenz von Politiklehrkräften basiert auf kompetenztheoretischen sowie kognitionspsychologischen Grundlagen und rekurriert explizit auf das generische Strukturmodell der COACTIV-Studie (Baumert & Kunter, 2006; Kunter et al., 2011) sowie die Konzeptualisierung der MT21-Studie (Blömeke et al., 2008) (Weschenfelder, 2014a, S. 101). Zum besseren Verständnis der PKP-Modellierungen sollen daher auch diese beiden Strukturmodelle vorgestellt werden.

Das COACTIV-Forschungsprogramm untersucht, „über welche personalen Voraussetzungen Lehrkräfte verfügen müssen, um ihren Beruf dauerhaft erfolgreich auszuüben“ (Baumert et al., 2011, S. 8). Hierzu wird ein Strukturmodell professioneller Lehrerkompetenzen vorgelegt, das die Bestandteile derjenigen Kompetenzen darstellen soll, die notwendig sind, um „Unterricht planen, inszenieren und interaktiv gestalten zu können“ (ebd.) (Abbildung 12).

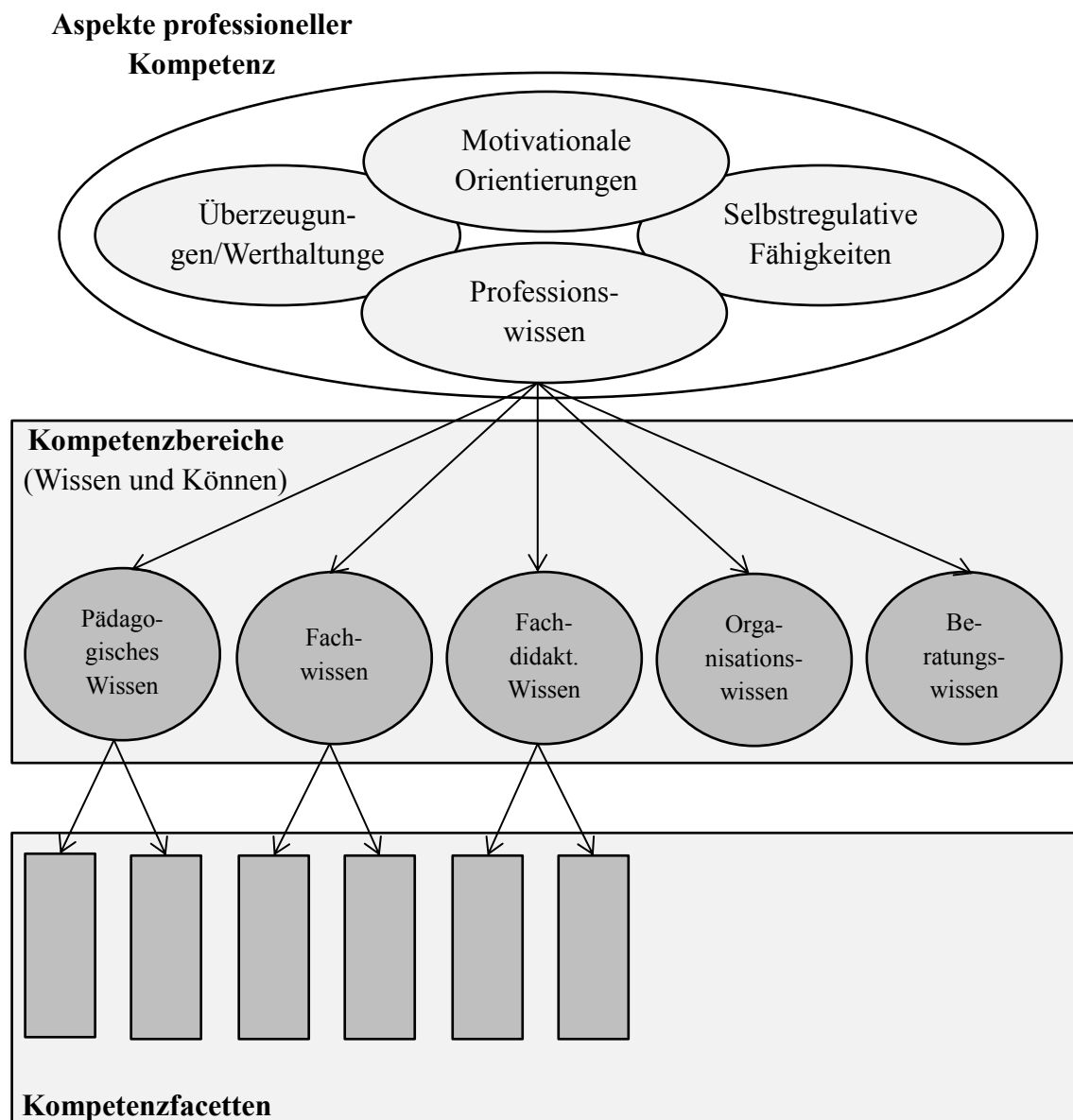


Abbildung 12: Allgemeines Kompetenzstrukturmodell professioneller Handlungskompetenzen (eigene Darstellung in Anlehnung an Baumert & Kunter, 2006, S. 482; Baumert & Kunter, 2011a, S. 32)

Der Begriff der professionellen Handlungskompetenz in COACTIV wird in Anlehnung an einen kognitionspsychologisch geprägten Kompetenzbegriff verwendet, umfasst also neben kognitiven Aspekten auch Merkmale motivationaler und selbstregulativer Eigenschaften. Im Sinne des Modells entsteht „professionelles Handeln aus dem Zusammenspiel von ... spezifischem, erfahrungsgesättigten deklarativen und prozeduralen Wissen (Kompetenzen im en-

geren Sinne: Wissen und Können); ... professionellen Werten, Überzeugungen, subjektiven Theorien, normativen Präferenzen und Zielen; ... motivationalen Orientierungen sowie ... Fähigkeiten der professionellen Selbstregulation [Umformatierung v. Verf.]“ (Baumert & Kunter, 2011a, S. 33). Das professionelle Wissen der Lehrkräfte ist im COACTIV-Modell im Wesentlichen entlang der Wissenstaxonomie von Shulman (1986, 1987) gegliedert, wobei von den dargestellten Wissensbereichen das Fachwissen, das fachdidaktische Wissen und das pädagogische Wissen als die zentralen Kategorien angesehen werden (Krauss et al., 2008b, S. 226; vgl. auch Kapitel 2.2).

Wie bereits in Tabelle 2 angerissen, wird das für das unterrichtliche Handeln als besonders wichtig erachtete fachdidaktische Wissen in COACTIV in enger Anlehnung an kognitionspsychologische Grundlagen in diesen drei Dimensionen unterschieden (Baumert & Kunter, 2011a):

- „Wissen über das didaktische und diagnostische Potenzial, die kognitiven Anforderungen und impliziten Wissensvoraussetzungen von Aufgaben, ihre didaktische Sequenzierung und die langfristige curriculare Anordnung von Stoffen,
- Wissen über Schülervorstellungen (Fehlkonzeptionen, typische Fehler, Strategien) und Diagnostik von Schülerwissen und Verständnisprozessen,
- Wissen über multiple Repräsentations- und Erklärungsmöglichkeiten.“ (S. 37f.)

Zusammengefasst beschreibt dieser Kompetenzbereich „Fähigkeiten theoretisch als berufs- und situationsspezifische erlernbare Facette ... *fachdidaktischen Wissens und Könnens*, wenn bei auftretenden Verständnisschwierigkeiten gezielt und einfühlsam konstruktive Hilfe gewährt werden muss“ (Baumert et al., 2011, S. 11).

Auch der zweite theoretische Bezugsrahmen für das PKP-Forschungsprogramm, die MT21-Studie (Blömeke et al., 2008) sowie die Folgestudien TEDS-M (Blömeke et al., 2010a; Blömeke et al., 2010b) und TEDS-LT (Blömeke et al., 2011; Blömeke et al., 2013), orientiert sich „an der kognitionspsychologischen Forschung, insbesondere der Expertiseforschung, und adaptierte diese für Lehrkräfte“ (Blömeke, 2013a, S. 7). Bezug genommen wird – ebenso wie in der theoretischen Rahmung des COACTIV-Forschungsprogramms – auf die Taxonomie des Lehrerwissens bei Shulman (1986, 1987). Für die Konzeptualisierung der TEDS-Studien wird betont, dass es „bedeutsam ist, dass professionelle Kompetenz von Lehrkräften über deklaratives Faktenwissen hinausgeht“ (Blömeke, 2013a, S. 11). Wie dem COACTIV-Programm lässt sich den TEDS-Studien hinsichtlich der Frage nach der

Handlungsvalidität des operationalisierten und erhobenen Wissens eine Transformationsposition (Vogelsang & Reinhold, 2013, S. 108) zuordnen; die theoretischen Konzeptualisierungen beider Forschungsprogramme betonen übereinstimmend unter Rückbezug auf Weinert (2001), dass vorhandenes Wissen als ‚Verfügenkönnen‘ von Leistungen interpretiert werden kann – oder in einer Formulierung von Blömeke (2013a):

„Bevor von Professionalität gesprochen werden kann, muss also nicht nur eine kognitive Integration von Wissensdimensionen erfolgen. Diese müssen auch mit spezifischen Erfahrungen und Vorstellungen verbunden sein, bevor sie wirksam werden können. Eine solche mehrdimensionale Konzeptualisierung der professionellen Kompetenzen angehender und praktizierender Lehrkräfte ist an die übergreifende Kompetenzdefinition von Weinert (1999) anschlussfähig. Die Konzeptualisierung ist damit auf Leistungsdisposition ausgerichtet, seien sie kognitiver oder affektiv-motivationaler Art. Die unterrichtliche Performanz, die von zahlreichen weiteren Kontextfaktoren beeinflusst wird, bleibt zunächst ausgespart.“ (S. 11)

Bei der Entwicklung eines domänenspezifischen Strukturmodells im PKP-Programm (Oberle et al., 2012) wird auf die theoretischen Grundlagen dieser Forschungsprogramme Bezug genommen, um „Kompetenzen [zu bestimmen], die für erfolgreiches Unterrichten (vgl. auch Krauss et al. 2004, Tenorth, 2006; Blömeke, 2003; Terhart, 2000) in Politik nötig sind“ (Weschenfelder, 2014a, S. 101) und dafür, „Struktur, Ausprägungen und Zusammenhänge professioneller Lehrerkompetenz für das Fach in einem Modell zu beschreiben“ (Weißeno et al., 2013, S. 188f.). Die Darstellung des Modells ist im Wesentlichen vergleichbar mit den generischen Strukturmodellen bei Baumert und Kunter (2006, 2011a) sowie Blömeke et al. (2008) (Abbildung 13).

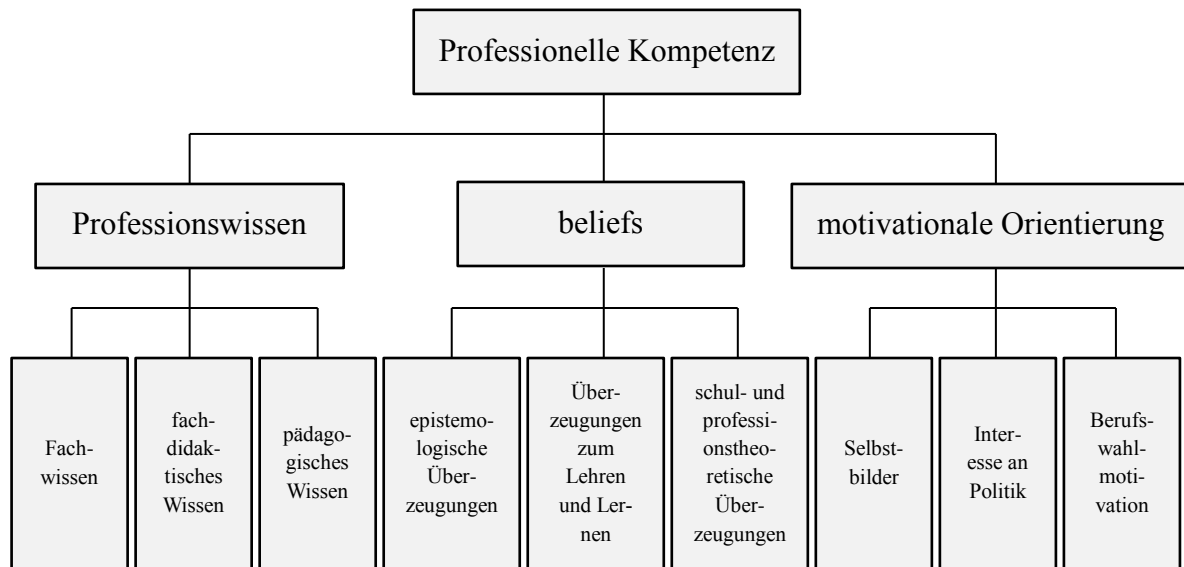


Abbildung 13: Modell professioneller Handlungskompetenz von Politiklehrkräften der PKP-Studie (eigene Darstellung nach Weißenö et al., 2013, S. 189)

Im PKP-Modell spiegelt „[k]ompetentes Handeln ... die Verknüpfung kognitiver und motivationaler Kompetenzen wider und zeigt sich in der Verbindung von Wissen und Können (vgl. Klieme & Hartig, 2007, S. 19)“ (Weschenfelder, 2014a, S. 101f.). Für die Unterscheidung von Wissen und Können wird das Modell von Neuweg (2014, S. 585) (Abbildung 4) herangezogen und es wird „die Relevanz einer kategorialen Trennung [betont], die auch in politikdidaktischen Kompetenzmodellen bei der Kompetenzbestimmung und -analyse durchzuhalten ist“ (Weschenfelder, 2014a, S. 77), weshalb es sich in diesem Sinne bei der professionellen Kompetenz von Politiklehrkräften „um ein Gemisch aus verschiedenen Facetten des Denkens und Handelns [handelt], die situationale Ausprägungen annehmen und sich in Denken und Mustererkennung auflösen“ (ebd.), was Herausforderungen nicht nur an die theoretische Konzeptualisierungen von Kompetenzstrukturen, sondern auch an die forschungspraktischen Zugänge zur Erfassung professionellen Wissens und Könnens stelle (ebd.).

Die Anmerkung von Vogelsang und Reinhold (2013), dass beim Wissenstest des COACTIV-Programms „[e]ine direkte Beobachtung des Unterrichtshandelns ... allerdings nicht in die Daten ein[floss]“ (S. 121) und die Handlungsvalidität daher kritisch gesehen werden müsse, kann auch auf das PKP-Forschungsprogramm bezogen werden. Es kann außerdem angemerkt werden, dass für die Entwicklung von Test-Items mit dem Bezug auf den Entwurf von

Batt et al. (2007; zitiert nach Weschenfelder, 2014a, S. 109) zwar ein theoretischer Referenzrahmen gewählt wurde, dieser jedoch trotz seines Anspruchs auf Domänenspezifität in den Ausformulierungen der Facetten politikdidaktischen Wissens weiterhin in Teilen fächergenerell erscheint (vgl. Kapitel 2.2.1). Das in diesem Zusammenhang nochmals zu betonende Problem der weiterhin teilweise ausstehenden empirischen Forschungen, die einen entsprechenden Begründungszusammenhang für die ‚inhaltliche Füllung‘ der Konzeptualisierung und Operationalisierung liefern könnten, wird auch in der PKP-Studie erkannt (Weschenfelder, 2014a, S. 74).

Es lässt sich zusammenfassen, dass die PKP-Modellierung (Oberle et al., 2012) als bislang wissenschaftstheoretisch elaboriertester Ansatz angesehen werden kann (vgl. auch Manzel, 2013, S. 208f.). Bei der Modellierung und Operationalisierung professioneller Handlungskompetenzen von Politiklehrkräften werden kompetenztheoretisch und kognitionspsychologisch begründete Referenzpunkte für die Modellierung ausgewiesen. Die besonderen Herausforderungen der Operationalisierung und Erhebung handlungsnaher Aspekte professioneller Kompetenzen werden auch im Rahmen der Entwicklung der PKP-Studie hervorgehoben; hiermit verbundene Forschungsbedarfe werden betont (Weschenfelder, 2014a, S. 285).

2.4 Politikdidaktische Überlegungen zur weiteren Modell- und Instrumentenentwicklung

Der Überblick über den aktuellen Theoriestand der Politikdidaktik zum Professionswissen mit besonderem Fokus auf dem fachdidaktischen Wissen von Politiklehrkräften lässt den Schluss zu, dass „[d]ie Leistungen im politikdidaktischen Wissen ... weiter untersucht werden [müssen]“ (Weißeno et al., 2013, S. 200). Zur Weiterentwicklung vorhandener Instrumente der Modellierung und Diagnostik politikdidaktischer Kompetenz lassen sich auf Grundlage der theoretischen Exploration folgende Überlegungen zusammenfassen.

Erweiterte mehrdimensionale Modellierung des Professionswissens (Wissen und Können)

Eine an die Forschungstradition anderer Fächerdomänen anschlussfähige und theoretisch begründete Modellierung der Struktur der Kompetenz von (angehenden) Politiklehrkräften liegt vor (Oberle et al., 2012; Oberle, Weschenfelder & Weißeno, 2013; Oberle,

Weschenfelder & Weißeno, 2014; Weißeno et al., 2013; Weschenfelder, 2014a; Weschenfelder, Weißeno & Oberle, 2014) und bietet demnach ein anschlussfähiges übergeordnetes Kompetenzmodell. Erfolgt jedoch die Modellierung von Handlungskompetenzen entlang der Konzeptualisierung professioneller Lehrerkompetenzen im COACTIV-Forschungsprogramm als *Wissen und Können* (Baumert & Kunter, 2006, S. 481; Baumert & Kunter, 2011a, S. 33), muss geklärt werden, „welches explizite Wissen für welches Können wann und für wen hinreichend, notwendig oder zumindest hilfreich ist“ (Neuweg, 2014, S. 600), was auch Weschenfelder (2014a, S. 79) im Rahmen der Darstellung der Modellierung des PKP-Forschungsprogramms betont.

Da sich „[e]rgänzend zu den bisher dargelegten Konzeptualisierungen ... das Professionswissen auf mehrere Arten messen [lässt]“ (Tepner et al., 2012, S. 16f.), sind hierfür Konzeptualisierungen notwendig, die diese „Wissensarten“ (ebd.) oder auch „*Wissenskomponenten*“ (Schmelzing, 2010, S. 29) als weiteren Aspekt der Modellierung adäquat abbilden (Klieme et al., 2007, S. 13). Entsprechende Modellierungen wurden bereits vorgestellt und es wird im weiteren Verlauf dieser Arbeit vorgeschlagen, die Modelle des Professionswissens bei ProwiN (Tepner et al., 2012) sowie des fachdidaktischen Wissens bei Schmelzing (2010) als Grundlagen für die weiterführende Ausdifferenzierung des Strukturmodells politikdidaktischer Kompetenzen von (angehenden) Lehrkräften und die domänenspezifische Instrumentenweiterentwicklung heranzuziehen.

Spezifizierung der Inhaltsbereiche

Eine um die Dimension *Inhaltsbereiche* erweiterte mehrdimensionale Modellierung des Professionswissens von Politiklehrkräften bietet darüber hinaus die Möglichkeit, „die fachlichen *Inhalte* (Shulman, 1987; Shulman, 1986; Veal & Kubasko, 2003), auf die sich die fachdidaktischen Kenntnisse und Fertigkeiten der Lehrkräfte beziehen“ (ebd., S. 29) zu präzisieren. Da sich das fachdidaktische Wissen und Können einer Lehrkraft entsprechend der Taxonomie von Shulman (1986, 1987) nur in Bezug auf die unterrichtlichen Inhalte definieren lässt (Krauss et al., 2008b, S. 227) und in anderen Domänen davon ausgegangen wird, dass es keine Hinweise auf „inhaltsunabhängige fachdidaktische Wissensbereiche“ (Gramzow et al., 2013, S. 13) gibt (vgl. auch Tepner et al., 2012, S. 9), werden in den nachfolgenden Modellierungsvorschlägen ebenfalls Inhaltsbereiche konkretisiert. Eine Anschlussfähigkeit an das

Kompetenzstrukturmodell des PKP-Forschungsprogramms ist weiterhin gegeben, da auch in diesem bei der Item-Entwicklung Schwerpunkte „auf den Bereichen politisches System der Bundesrepublik Deutschland, Politische Systeme im Vergleich und Internationale Beziehungen [liegen]“ (Weschenfelder, 2014a, S. 167).

Neben den damit verbundenen Vorteilen für die Operationalisierung der Konstrukte des politikdidaktischen Wissens ist gleichzeitig ein – wenn auch zunächst probeweiser – Ansatz aufgezeigt, dem Problem einer unter Umständen unklaren Domänenspezifität bei weiterführenden Konzeptualisierungsansätzen zur Modellierung und Erfassung sozialwissenschaftlicher Kompetenzen von Lehrkräften zu begegnen.

Erhöhung des Auflösungsgrades der Modellierung

Eng mit der Spezifizierung der inhaltlichen Dimension der Modellierung verbunden ist die Möglichkeit einer präzisen Operationalisierung der Konstrukte. Krauss et al. (2008b) weisen im Hinblick auf die Testkonstruktion zu den beiden zentralen Kompetenzfacetten des fachdidaktischen Wissens darauf hin, dass diese „im Wesentlichen Shulmans Konzeptualisierung folgen kann (die ‚nur‘ noch mit Fachinhalt gefüllt werden muss)“ (S. 233). Hiermit ist ein Verfahren angesprochen, das „einer deduktiven Konstruktionsweise [folgt], indem ein theoretisches Konstrukt beschrieben und aus diesem mögliche beobachtbare Indikatoren abgeleitet werden“ (Weschenfelder, 2014a, S. 163). Dies ist im späteren Verlauf bezogen auf den in der vorliegenden Arbeit interessierenden Inhaltsbereich zu leisten. Hierzu werden in Kapitel 3.2 in einer Kombination aus induktiven und „normativ-deduktiven Verfahren“ (Schaper, 2009) diejenigen Facetten der politikdidaktischen Kompetenz konzeptualisiert, die sich auf die in dieser Untersuchung relevante Inhaltsdimension beziehen.

Erprobung videobasierter Verfahren zur Diagnostik politikdidaktischer Kompetenzen

Aus den dargestellten Ansätzen einer Erweiterung der vorliegenden Modellierungen und Operationalisierungen professioneller Handlungskompetenzen von (angehenden) Politiklehrkräften ergeben sich erweiterte Anforderungen an eine angemessene Testung. Wie bereits problematisiert, eignen sich traditionelle schriftliche Tests zwar für die Erfassung deklarativen Professionswissens, jedoch – je nach theoretischer Sichtweise – nur bedingt oder gar

nicht, um handlungsbezogene Aspekte professioneller Handlungskompetenzen zu erfassen. Denn „[o]b eine Lehrperson handlungskompetent bezogen auf das Unterrichten ist, zeigt sich darin, dass diese Person in der Lage ist, diese Anforderung zu bewältigen“ (Vogelsang & Reinhold, 2013, S. 106). Es konnte dabei in Kapitel 2.3.1.1 zwar dargelegt werden, dass aus theoretischer Sicht die beiden theoretischen Konstrukte Kompetenz und Performanz unterschiedliche Gegenstandsbereiche beschreiben, jedoch wurde gleichzeitig darauf hingewiesen, dass Kompetenz „[i]m engeren Sinne ... nur über die Betrachtung der *Performanz*, also beobachtbares Verhalten, dem sich Qualitäten zuordnen lassen, feststellbar [ist] (vgl. Klieme & Hartig, 2007)“ (Vogelsang & Reinhold, 2013, S. 106), da Kompetenzen latente Persönlichkeitsmerkmale darstellen, die nicht direkten Testverfahren zugänglich sind. Dies bringt neue Herausforderungen für die Ansätze zur Erfassung politikdidaktischer Kompetenzen mit sich.

Testverfahren, in denen die Kompetenzen von Lehrpersonen beispielsweise in Form von Selbst- und/oder Fremdbeurteilungen diagnostiziert werden, sind möglich, müssen aber schon im Hinblick auf Aspekte wie verzerrte Selbst- und Fremdwahrnehmung und soziale Erwünschtheit problematisiert werden (Maag Merki & Werner, 2014, S. 752). Stattdessen werden computerbasierte Verfahren vorgeschlagen, die den Probanden im Format des video-cued testing Stimuli (Videovignetten) mit authentischen und anforderungsreichen bzw. problemhaltigen Unterrichtssituationen präsentieren, auf die sie reagieren müssen, als wenn sie im Unterricht wären. Zwar gilt es bei computerbasierten Tests, nicht unwesentliche Probleme für die Forschung zu bewältigen, jedoch wird solchen Erhebungsverfahren das Potenzial zugeschrieben, die Komplexität des Lehrerberufs angemessener darzustellen, als es durch schriftliche Tests möglich ist (Knievel et al., 2015). In Anlehnung an das computerbasierte Verfahren bei Lindmeier (2011) wird im weiteren Verlauf der vorliegenden Arbeit die Übertragung dieses Verfahrens zur Diagnostik ausgewählter Aspekte professioneller Kompetenzen von (angehenden) Politiklehrkräften erprobt und es wird kritisch diskutiert, welche Implikationen sich aus der explorativen Studie für zukünftige Erhebungsverfahren in diesem Bereich ergeben können.

Berücksichtigung qualitativer Aspekte politikdidaktischer Kompetenzen

Zur Beurteilung politikdidaktischer Kompetenzen sind qualitative Gütemaßstäbe notwendig. Können beispielsweise bei Multiple-Choice-Aufgaben zum politikdidaktischen Wissen (vgl. auch Abbildung 5) Bewertungen noch anhand dichotomer Kategorien erfolgen (richtig/falsch), ist dies bei halb offenen Antwortformaten bereits schwieriger. Für die Beurteilung sprachlicher Rückmeldungen von Probanden beim video-cued testing, die durch ein erheblich höheres Maß an Komplexität und Deutungsvielfalt gekennzeichnet sind, erscheint schließlich die Berücksichtigung erweiterter qualitativer Gütemaßstäbe unabdingbar.

In der Übersicht über qualitative Merkmale von Unterricht, die mit der Expertise der Lehrkräfte zusammenhängen, wurden ausgewählte Aspekte aus einer domänenspezifischen Perspektive herausgearbeitet, wobei deutlich geworden ist, dass es eine Aufgabe auch für die Politikdidaktik ist, „Unterrichtsmerkmale fachlich und fachdidaktisch auszudifferenzieren“ (Klieme & Rakoczy, 2008, S. 235). Die von Klieme, Lipowsky, Rakoczy und Ratzka (2006) vorgeschlagenen Basisdimensionen zeigen sich dabei im Kern kompatibel zu bislang vorliegenden politikdidaktischen Überlegungen. Der wissenschaftstheoretische Status einiger politikdidaktischer Ansätze muss wegen teilweise fehlender Bezugnahme auf entsprechende übergeordnete Theorien sowie noch ausstehender empirischer Befunde kritisch reflektiert werden. Dennoch sollen diese bei der weiteren Instrumentenentwicklung – hier insbesondere bei der Ausgestaltung des Ratingmanuals zur Beurteilung der offenen Antwortkategorien – berücksichtigt werden, um einen qualitativen Referenzrahmen festlegen zu können.

3 Erweiterte Modellierung professionellen Wissens und Könnens von Politiklehrkräften

Auf Grundlage der bisherigen theoretischen Überlegungen soll in diesem Kapitel zunächst ein mehrdimensionales Modell des Professionswissens von Politiklehrkräften vorgeschlagen werden (Kapitel 3.1), bevor anschließend ein heuristisches mehrdimensionales Modell des politikdidaktischen Wissens beschrieben wird, das für die Entwicklung von Test-Items in dieser Studie herangezogen wird (Kapitel 3.2).

3.1 Ein erweitertes mehrdimensionales Modell des Professionswissens von Politiklehrkräften

Mit der mehrdimensionalen Konzeptualisierung des Professionswissens von Tepner et al. (2012) liegt ein Modell vor, das die Strukturdimensionen *Wissensarten* und *Inhaltsbereiche* implementiert. Dabei ist das Modell, obwohl es „primär der Konstruktion von Testitems zur Erfassung des Professionswissens in den drei Naturwissenschaften [dient], ... aber gleichzeitig auch eine relative Vergleichbarkeit der Ergebnisse zu den drei Dimensionen des Professionswissens (CK, PK, PCK) zwischen den beteiligten naturwissenschaftlichen Fächern ermöglichen [soll]“ (ebd., S. 18), trotz seiner domänenspezifischen Ausformulierung ausreichend ‚offen‘ und im Wesentlichen anschlussfähig an die fachspezifische Modellierung professioneller Handlungskompetenzen von Politiklehrkräften in der PKP-Studie (Oberle et al., 2012). Es sind jedoch Modifikationen im Sinne der ‚inhaltlichen Füllung‘ der erweiterten Strukturdimensionen notwendig. Unter Berücksichtigung der Theorien zu den Arten des Wissens sowie der in Kapitel 2.2.2 erläuterten und im Folgenden angenommenen Inhaltsspezifität des fachlich-fachdidaktischen Wissens von Lehrkräften lässt sich ein domänenspezifisches mehrdimensionales Modell des Professionswissen von Politiklehrkräften zusammenfassen (Abbildung 14).

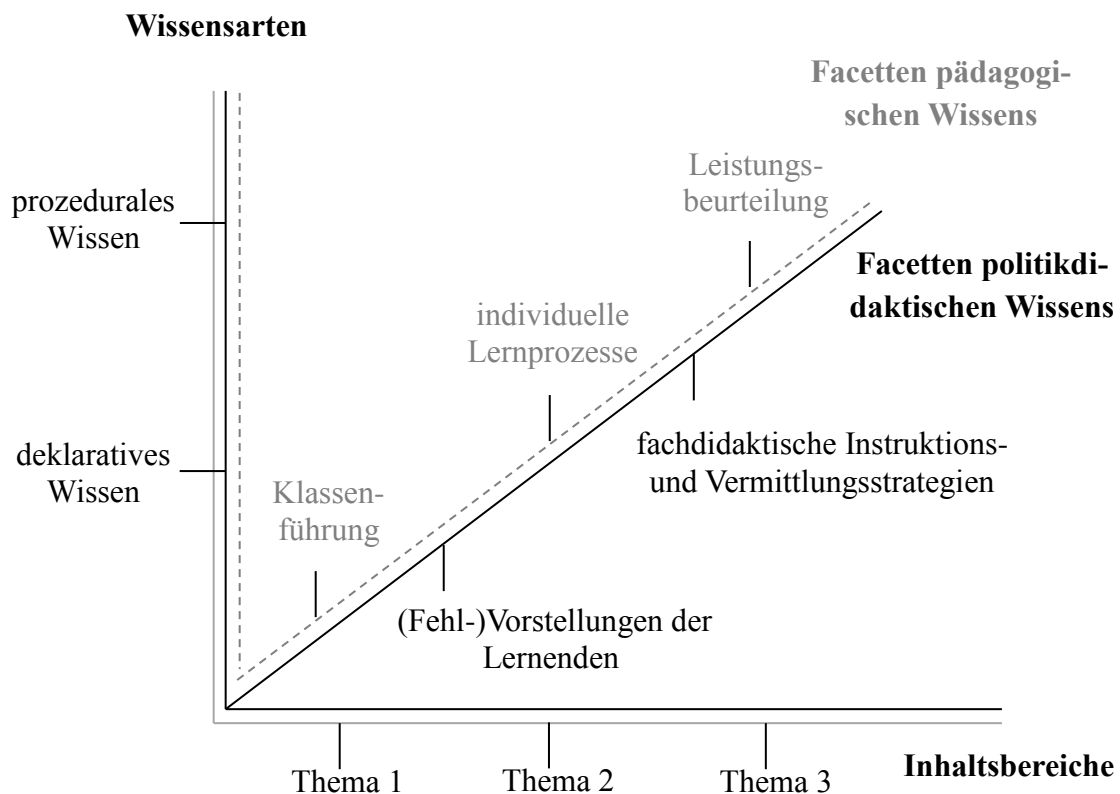


Abbildung 14: Mehrdimensionales Modell des Professionswissens von Politiklehrkräften (in Anlehnung an Tepner et al., 2012, S. 19 sowie ergänzend Baumert & Kunter, 2011a; Krauss et al., 2008b; Schmelzing, 2010; Weißeno et al., 2013)

Das Modell synthetisiert die im bisherigen theoretischen Rahmen diskutierten Ansätze zur Konzeptualisierung und Modellierung des Professionswissens unter einer domänenspezifischen Perspektive. Im strukturellen Zusammenhang zu den im Verlauf dieser Arbeit referierten allgemeinen sowie fachspezifischen Kompetenzstrukturmodellen (Baumert & Kunter, 2011a; Weißeno et al., 2013; vgl. auch Abbildung 12 und Abbildung 13) zeigt das Modell einen *Aspekt professioneller Kompetenz* (Professionswissen von Politiklehrkräften) in seinen zentralen *Kompetenzbereichen* bzw. *Kompetenzdimensionen* (Fachwissen, fachdidaktisches Wissen und pädagogisches Wissen) sowie assoziierten *Kompetenzfacetten* (im politikdidaktischen Wissen die Facetten des Wissens über fachdidaktische Instruktions- und Vermittlungsstrategien sowie des Wissens über [Fehl-]Vorstellungen der Lernenden) auf. Der Modellierung liegt dabei ein kognitionspsychologisches Kompetenzverständnis zugrunde,

das diese „als *kontextspezifische kognitive Leistungsdispositionen* [definiert], die sich funktional auf Situationen und Anforderungen in bestimmten *Domänen* beziehen“ (Klieme & Leutner, 2006, S. 879), wobei „[d]ie Begrenzung auf kognitive Dispositionen bedeutet, dass motivationale und affektive Variablen im Bildungsgeschehen separat erfasst werden und Zusammenhänge zwischen Kompetenzen und diesen nicht-kognitiven Variablen einer expliziten empirischen Untersuchung zugänglich sind“ (Klieme et al., 2007, S. 7f.). In Anlehnung an die Modellierung der professionellen Handlungskompetenzen von Politiklehrkräften im PKP-Forschungsprogramm werden die Dimensionen und Facetten des Professionswissens somit unter der Prämisse einer *fokussierten Sichtweise* dargestellt (Weschenfelder, 2014a, S. 102). Es wird im Modell infolgedessen das Hauptaugenmerk auf das professionelle Wissen und Können als „Kern der Professionalität“ (Baumert & Kunter, 2006, S. 481) gelegt, was jedoch ausdrücklich nicht bedeuten soll, dass weitere Aspekte professioneller Handlungskompetenzen von Lehrpersonen wie Fähigkeiten der Selbstregulation sowie professionelle Überzeugungen (Baumert & Kunter, 2011a, S. 41ff.; Weschenfelder, 2014a, S. 110ff.) nicht als wichtig erachtet werden. Im Gegenteil sollte in zukünftigen Forschungen zum Beispiel der Zusammenhang von Selbstkonzepten als weiterer Einflussvariable für die Gestaltung des Unterrichts (Weschenfelder, 2014a, S. 131) zu den hier inhaltsspezifisch konzeptualisierten Kompetenzfacetten untersucht werden; der heuristische Charakter des Modells bietet dabei Möglichkeiten zur weiteren Ausdifferenzierung an.

Die Konzeptualisierung des *Fachwissens* von Politiklehrkräften erfolgt grundständig in Anlehnung an das Vorgehen zur Konstruktion von Test-Items zu dieser Kompetenzdimension im PKP-Forschungsprogramm (ebd., S. 162ff.) unter Rückgriff auf die Empfehlungen der Deutschen Vereinigung für Politische Wissenschaft (DVPW, 2003) sowie auf den Entwurf von Batt et al. (2007; zitiert nach Weschenfelder, 2014a, S. 109). Da empirische Befunde aus anderen Fächerdomänen wie der Mathematik darauf hindeuten, dass auch innerhalb eines Faches das Fachwissen von Lehrpersonen themenspezifisch unterschiedlich ausgeprägt sein kann (Hill et al., 2004), wird der Kompetenzbereich des Fachwissens im Modell wie bei Schmelzing (2010) und Tepner et al. (2012) inhaltsbezogen dargestellt. Qualitative Studien in der Mathematik deuten darauf hin, dass „eine direkte Abhängigkeit der im Unterricht verfügbaren Repertoires an Lehrstrategien und Erklärungsansätzen (fachdidaktisches Wissen) vom konzeptuellen Verständnis des Faches (Fachwissen) angenommen werden [kann]

(Baumert et al., 2010)“ (ebd., S. 20). Zwar können diese Befunde nicht unmittelbar auf die Politikdidaktik übertragen werden, dennoch sollen sie tentativ als Hinweise für die Modellierung des Fachwissens im Rahmen dieser Studie dienen. Vor dem Hintergrund fehlender empirischer Hinweise zum Beispiel zur Wirksamkeit fachdidaktischer Methoden und Prinzipien (Manzel & Neumann, 2015, S. 76f.) sowie in der Annahme, dass diese fachlich konkretisiert werden müssen (Sander, 2005, S. 29), wird dabei davon ausgegangen, dass eine inhaltspezifische Ausdifferenzierung des Fachwissens auch bei einer ‚organischeren‘ Operationalisierung des politikdidaktischen Wissens helfen kann. Es wird dabei gemäß dem transformativen Modell (Gess-Newsome, 1999) angenommen, dass es sich bei den einzelnen Kompetenzbereichen um voneinander trennbare Dimensionen handelt, was auch eine separate Erhebung des fachdidaktischen Wissens sowie des Fachwissens ermöglicht; somit ist eine Anschlussfähigkeit dieser Modelldarstellung des Fachwissens von Politiklehrkräften an die PKP-Modellierung (Weißeno et al., 2013, S. 189) weiterhin gegeben.

Die Konzeptualisierung der *Wissensarten* erfolgt entlang der in Kapitel 2.1.2 dargestellten kognitionspsychologischen Annahmen und unter Berücksichtigung der Rezeption dieser Grundlagen in den COACTIV- und PKP-Forschungsprogrammen (Baumert und Kunter, 2006, S. 481; Weschenfelder, 2014a, S. 51f.) in Form der analytischen Trennung des Wissens in deklaratives Wissen und prozedurales Wissen. Da in einigen Untersuchungen weitere Ausdifferenzierungen der Arten des Professionswissen vorgenommen werden (vgl. u. a. Schmelzing et al., 2010; Tepner et al., 2012), erfolgt hierbei die Fokussierung auf diese beiden Wissensarten zunächst aus einer forschungspragmatischen Perspektive; die Heuristik des Modells erlaubt die Erweiterung dieser Dimension in Folgestudien.

Der Kompetenzbereich des *pädagogischen Wissens* schließlich fußt in seinem inhaltlichen Verständnis zum Zwecke der Anschlussfähigkeit ebenfalls auf der Konzeptualisierung dieser Kompetenzdimension im Rahmen der COACTIV-Studie (Baumert & Kunter, 2011a, S. 38f.; Voss & Kunter, 2011) sowie der Nachfolgestudien, obgleich sie in der vorliegenden Arbeit zunächst unberücksichtigt bleiben soll.

3.2 Konzeptualisierung des politikdidaktischen Wissens und Könnens von Lehrkräften

Die mehrdimensionale Modellierung des Professionswissens erlaubt eine differenzierte Darstellung des in diese Arbeit im Mittelpunkt stehenden politikdidaktischen Wissens und Könnens in Form eines heuristischen mehrdimensionalen Strukturmodells (Abbildung 15).

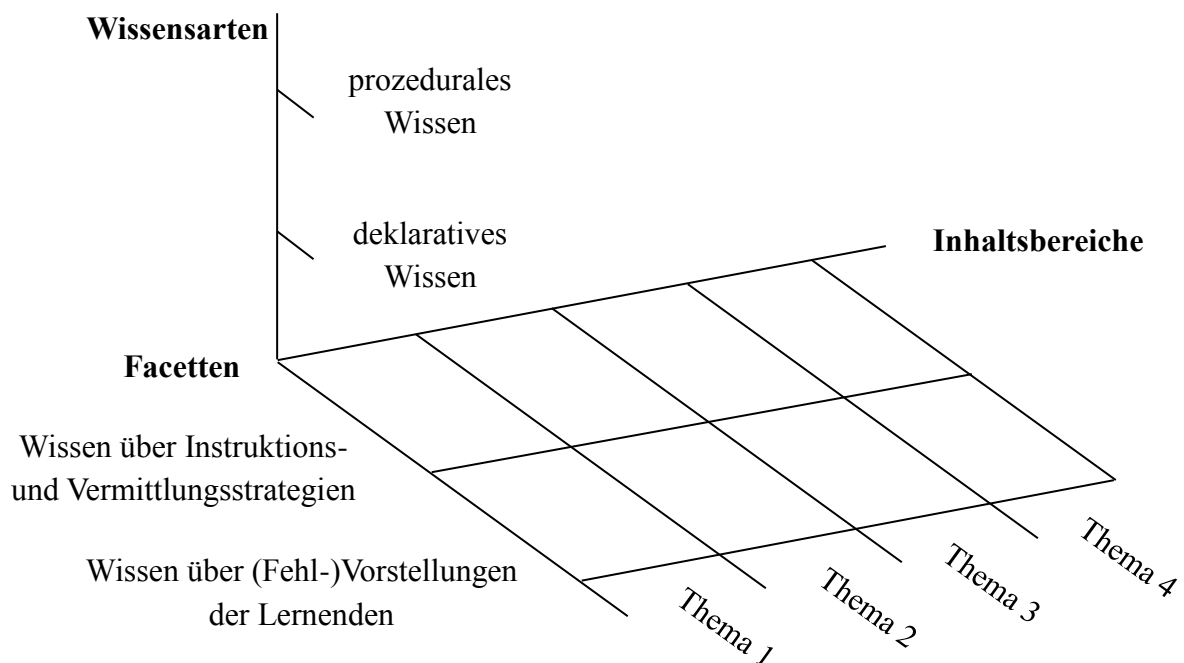


Abbildung 15: Heuristisches mehrdimensionales Strukturmodell des politikdidaktischen Wissens (in Anlehnung an Schmelzing, 2010; Tepner et al., 2012; Weißeno et al., 2013 sowie ergänzend Faissner et al., o. J., S. 7)

Dieses Modell dient in Anlehnung an die Modellierung des Professionswissens bei Tepner et al. (2012, S. 19) dazu, einen theoretisch fundierten Begründungszusammenhang zur (Weiter-)Entwicklung von Test-Items zum politikdidaktischen Wissen herstellen zu können, und soll im Folgenden in seinen zugrunde liegenden Konzeptualisierungen näher erläutert werden. Auf Grundlage der Empfehlung von Schaper (2009) wird eine Kombination aus einem normativ-deduktiven Vorgehen („Ableitung“ der Kompetenzstrukturdimensionen und deren Konzeptualisierung aus bestehenden Kompetenzmodellierungen sowie – sofern vorhanden –

empirischen Befunden) und einem deduktiven Vorgehen (Modellierung unter Berücksichtigung der beruflichen Anforderungssituationen) angewendet (Frey, 2014, S. 714).

3.2.1 Konzeptualisierung der Strukturdimension *Facetten*

In dieser Strukturdimension wird im Modell auf Grundlage der theoretischen Übersicht in Kapitel 2.2.1 davon ausgegangen, dass das *Wissen über Instruktions- und Vermittlungsstrategien* sowie das *Wissen über (Fehl-)Vorstellungen der Lernenden* die beiden zentralen Facetten des politikdidaktischen Wissens sind. Unter der Kompetenzfacette des Wissens über Instruktions- und Vermittlungsstrategien wird in Anlehnung an Baumert und Kunter (2011a) das „Wissen über multiple Repräsentations- und Erklärungsmöglichkeiten“ (S. 38) verstanden. Unter Berücksichtigung des Unterrichtsqualitätsmerkmals der Angebotsvielfalt (Helmke, 2014, S. 263ff.) wird daher angenommen, dass hierunter nicht *eine bestimmte* Methode für den Politikunterricht zu verstehen ist, oder wie es Brophy (2000) ausdrückt: „No single teaching method [...] can be the method of choice for all occasions. An optimal programme will feature a mixture of instructional methods and learning activities“ (S. 6, zitiert nach Helmke, 2014, S. 265). Nicht zuletzt wegen empirischer Befunde zu den lernförderlichen Effekten direkter Instruktion (*direct instruction*) (Wiechmann, 2011) ist zu betonen, dass sich eine adäquate Angebotsvielfalt im Hinblick auf Unterrichtsmethoden sowohl auf ein Lernangebot in Form offener Lernarrangements (Niggli, 2000) bezieht als auch auf Angebote, die ein höheres Ausmaß an Lenkung – im Sinne der Klassifikation von Meyer (2011, S. 78ff.) zum Beispiel auch Lehrgänge – beinhalten. So heben auch Manzel und Neumann (2015, S. 77) den Aspekt angemessener Methodenvariation aus politikdidaktischer Sicht hervor. In Anlehnung an den im Rahmen der Konzeptualisierung des politikdidaktischen Wissens in der PKP-Studie (Oberle et al., 2012) herangezogenen Vorschlag von Batt et al. (2007, zitiert nach Weschenfelder, 2014a, S. 109) überspannt diese Facette „Methoden, Arbeitstechniken und Medien für den Unterricht“ (ebd.) ebenso wie „Konzepte didaktischer Analyse und Planung sowie Evaluation von Unterricht“ (ebd.). Unter Berücksichtigung des normativen Diskurses in der Politikdidaktik wird hierunter in dieser Arbeit auch das Wissen über fachdidaktische Prinzipien gefasst. Allerdings ist darauf hinzuweisen, dass dieses ausdrücklich als ‚Wissen über ...‘ verstanden wird, weshalb eine Abgrenzung zum Verständnis über die Funktion von fachdidaktischen Prinzipien wie beispielsweise bei Reinhardt (2009, 2011, S. 152, 2012a; vgl. auch Kapitel 2.3.4) betont wird.

Unter der Kompetenzfacette des Wissens über (Fehl-)Vorstellungen der Lernenden wird in Anlehnung an die Konzeptualisierung bei Krauss et al. (2008b) im Rahmen des COACTIV-Forschungsprogramms das Wissen über „*typische Schülerfehler und -schwierigkeiten*“ (S. 234) verstanden. Diese Auffassung ist anschlussfähig an die Konzeptualisierung im PKP-Forschungsprogramm, in dem unter anderem das Wissen über „mentale Modelle und Misconceptions bei Lernenden“ (Batt et al., 2007; zitiert nach Weschenfelder, 2014a, S. 109) als Bestandteil dieser Kompetenzfacette angesehen wird (ebd.). Kenntnisse über korrekte und inkorrekte Schülervorstellungen (Präkonzepte) sowie über Fehlkonzepte gelten, wenn sie lernwirksam in die Unterrichtsplanung und -gestaltung eingebettet werden, als Prädiktor für hohen Lernerfolg (Born, 2007). Es ist dabei hervorzuheben, dass „Schülerfehler und typische Schwierigkeiten als eine didaktische Chance für verständnisvolles Lernen“ (Krauss et al., 2008b, S. 234) angesehen werden, da „Fehler ... ein natürlicher Bestandteil des Lernens [sind]“ (Helmke, 2014, S. 229), die „die Chance [eröffnen], einen Blick auf Missverständnisse und fehlgeschlagene Verstehensprozesse auf Seiten der Schüler wie auf Mängel in der Vermittlung des Unterrichtsangebotes werfen zu können“ (ebd.).

3.2.2 Konzeptualisierung der Strukturdimension *Inhaltsbereiche*

Diese Strukturdimension umfasst diejenigen Unterrichtsinhalte, anhand derer sich das politikdidaktische Wissen inhaltsbezogen abzeichnet (für das biologiedidaktische Wissen vgl. Schmelzing, 2010, S. 31). In Anlehnung an die Konzeptualisierung des Fachwissens im PKP-Forschungsprogramm wird es im Rahmen dieses Modells ebenfalls als notwendig angesehen, das Fachwissen und das fachdidaktische Wissen getrennt zu erfassen (Weschenfelder, 2014a, S. 107). Die im Modell aufgrund seines heuristischen Charakters nicht weiter spezifizierten Themen drücken dabei aus, dass sich die Facetten des politikdidaktischen Wissens innerhalb einer gewissen Bandbreite ausdifferenzieren lassen sollten, da sie sich auf *bestimmte* Inhaltsbereiche beziehen. So umfasst das Wissen über Instruktions- und Vermittlungsstrategien unter anderem Kenntnisse über unterschiedliche Methoden, Medien und didaktische Prinzipien, die der Lehrkraft das themenspezifische Erklären und Verständlichmachen ermöglichen (vgl. Kapitel 6.3.1 zur Operationalisierung in der vorliegenden Arbeit). Unter dem Wissen über (Fehl-)Vorstellung von Lernenden werden Kenntnisse zu

themenspezifischen Prä- und Fehlkonzepten²⁹ von Schülerinnen und Schülern gefasst, die dann als Ausgangspunkt für kognitive Aktivierung und fachlichen Konzeptaufbau durch eine gezielte Aufbereitung und Steuerung des Lernangebotes genutzt werden. Dass sich das Wissen von Politiklehrkräften über das typische (prä-)konzeptuelle Wissen der Schüler/-innen inhaltsspezifisch ausdifferenzieren lässt, kann auch mit Bezug auf das Modell der Politikkompetenz von Lernenden (Detjen et al., 2012, S: 29ff.) begründet werden. Die Autoren betonen, dass es „[z]um Initiieren des Konzeptaufbaus bei Schüler/-innen ... für Lehrende wichtig [ist], deren Präkonzepte diagnostizieren zu können“ (Weißeno et al., 2010, S. 50). Zu den unterschiedlichen Fachkonzepten der Politik (für eine Übersicht über die Basis- und Fachkonzepte vgl. ebd., S. 12) werden jeweils assoziierte Fehlvorstellung bzw. Fehlkonzepte dargestellt, die bei Lernenden vorhanden sein können.

3.2.3 Konzeptualisierung der Strukturdimension Wissensarten

Das *deklarative* politikdidaktische Wissen im Modell umfasst faktenbezogenes Wissen über die inhaltsrelevanten Instruktions- und Vermittlungsstrategien sowie (typischen) (Fehl-)Vorstellungen der Lernenden, insofern also „explizit-deklarative Kenntnisse über fachdidaktische Fakten, Begriffe und Zusammenhänge [*sic*] aber auch explizit-deklarative Kenntnisse über adäquate fachdidaktische Einsatzbedingungen (vgl. Polanyi, 1958; Ryle, 1949; Schelten, 2004)“ (Schmelzing, 2010, S. 30). Beispiele für ein solches deklaratives politikdidaktisches Wissen sind methodische Kenntnisse über die Schrittfolge der Karikaturenanalyse im Sinne politikdidaktischer Hermeneutik, Kenntnisse zu den (normativen und Kompetenz-)Zielen des Faches oder auch das Wissen über typische Verständnisschwierigkeiten der Lernenden im jeweiligen Inhaltsbereich.

Ergänzend zum deklarativen Wissen umfasst diese Strukturdimension das *prozedurale* politikdidaktische Wissen im Sinne politikdidaktischen Könnens als „mehrheitlich implizites und handlungsbezogenes Verfahrenswissen, automatisierte Handlungsrouniten sowie peri-aktionale, das heißt simultan zur Handlung vorliegende, Wahrnehmungs- und Verarbeitungsroutinen (Baumert et al., 2004; Baxter & Lederman, 1999; Park & Oliver, 2008; Tamir, 1988; van Driel et al., 1998; Weinert et al., 1990)“ (Schmelzing, 2010, S. 31). Dieses Wissen

²⁹ Zur Kritik an dem Begriff der *Fehlkonzepte* von Lernenden vgl. u. a. Autorengruppe Fachdidaktik (2011, S. 166ff.) und Lange (2011, S. 100).

dient den Lehrpersonen dazu, dass in einer unterrichtlichen Anforderungssituation „[b]estimmte Elemente der Problemsituation mit bestimmten Lösungsschritten verknüpft [werden]“ (Bromme, 1992/2014, S. 43)³⁰. Dass eine Lehrkraft unter Zeitdruck einen Fehler diagnostizieren „und auf diesen spontan fachdidaktisch wirksam ... reagieren“ (Schmelzing et al., 2010, S. 192) kann, beispielsweise durch eine erklärende Rückmeldung oder einen unterstützenden Impuls (Lindmeier, 2011; Lindmeier et al., 2012), ist hierfür ein Beispiel. Die Konzeptualisierung von reflexivem Wissen nach Schmelzing (2010) bzw. reflexiver Kompetenz nach Lindmeier et al. (2012) als „Fähigkeit, die Lehrkräfte auf Grundlage ihres Basiswissens benötigen, um die professionellen Anforderungen außerhalb des eigentlichen Unterrichtsprozesses zu bewältigen“ (S. 106), bleibt aufgrund des Fokus der Arbeit auf die Zusammenhänge zwischen theoretisch-formalem Wissen und handlungsbezogenen Wissen für die vorliegende Studie zunächst ausgespart. Es wird hierbei davon ausgegangen, dass „[d]er [unterrichtliche] Zeitdruck ... es der Lehrkraft eben nicht [erlaubt], reflexive Prozesse außerhalb der Handlung zu aktivieren (vgl. Schön 2002, reflection-in-action)“ (ebd.; für einen Ausblick auf mögliche Anchlusserhebungen vgl. auch Kapitel 8.2).

³⁰ Wie Bromme (1992/2014, S. 115ff.) aufzeigt, kann der Begriff des (*komplexen*) Problemlösens auch auf das Unterrichten übertragen werden, ohne den Anforderungsreichtum des Unterrichtsprozesses zu unterschätzen.

4 Empirische Evidenzen

Der Forschungsstand zu den professionellen Kompetenzen von Politiklehrkräften ist in der Politikdidaktik im Vergleich zu anderen Fächerdomänen weiterhin eher gering ausgeprägt (Weißeno et al., 2013, S. 187f.). Vor allem die Mathematikdidaktik sowie die naturwissenschaftlichen Didaktiken können auf einen umfangreicheren Forschungskontext blicken. Daher sollen nachfolgend auch ausgewählte empirische Befunde aus anderen Fächerdomänen betrachtet werden, sofern sie für die vorliegende Arbeit relevant sind.

Das COACTIV-Forschungsprogramm zeigt auf, dass sich das Fachwissen sowie das fachdidaktische Wissen voneinander unterscheiden lassen, „wobei bei höherem Expertisegrad eine zunehmende Verschmelzung zu beobachten ist“ (Kunter & Baumert, 2011, S. 347). Die Bedeutung des fachdidaktischen Wissens wird hervorgehoben, da „besseres Fachwissen ... weder einen direkten Einfluss auf die kognitive Aktivierung im Unterricht noch auf die individuelle Unterstützung beim Lernen [hat], die Lehrkräfte anbieten können, wenn Schwierigkeiten auftreten“ (Baumert & Kunter, 2011b, S. 183). Ein zentraler Befund von COACTIV ist, dass das fachdidaktische Wissen „auf Klassenebene knapp 40 % der Varianz der PISA-Mathematikleistung in Klasse 10 erklären kann“ (Krauss et al., 2008b, S. 251); es gilt als Prädiktor für die Lernleistungen der Schüler/-innen (ebd.; vgl. auch Neuweg, 2014, S. 591). Die Untersuchungen der sogenannten Michigan-Gruppe (Ball et al., 2005) konnten ebenfalls nachweisen, dass das „mathematical knowledge for teaching“ (S. 373) von Grundschullehrkräften, das Aspekte des Fachwissen sowie des fachdidaktischen Wissens umfasst, den Lernfortschritt vorhersagen kann.

Auch in der physikdidaktischen Forschung bestätigen empirische Befunde, dass es sich beim Professionswissen von Lehrkräften nicht um ein Gesamtkonstrukt handelt, sondern dass die einzelnen Kompetenzdimensionen voneinander trennbare Bereiche darstellen (Riese & Reinhold, 2012, S. 127), die jedoch in einem Bezug untereinander stehen dürften. Dem deklarativen physikdidaktischen Wissen kommt dabei „eine Mediatorfunktion in Bezug auf die anderen Bereiche des Professionswissens [zu]“ (Riese & Reinhold, 2010, S. 182) und das Fachwissen sowie das pädagogische Wissen werden als Voraussetzung für das physikdidaktische Wissen angesehen (ebd.). Deutliche Korrelationen zwischen dem mathematikbezogenen Fachwissen und dem mathematikdidaktischen Wissen angehender Lehrkräfte zeigen sich ebenso in der MT21-Studie (Blömeke et al., 2008). Und auch die Autoren des PKP-

Forschungsprogramms berichten für die professionelle Handlungskompetenz von Politiklehrkräften von ähnlichen positiven Zusammenhängen zwischen den Wissenskomponenten Fachwissen, fachdidaktisches Wissen und pädagogisches Wissen bei gleichzeitigen Hinweisen darauf, dass es sich bei den modellierten Komponenten nicht um ein gemeinsames Konstrukt handelt (Weißeno et al., 2013, S. 191). Demnach bildet „[e]in dreidimensionales Modell, das politikdidaktisches Wissen in Wissen um normative fachdidaktische Konzeptionen und Diskurse und in ein stärker unterrichtsbezogenes Wissen um Fehlkonzepte, Fördermöglichkeiten und den Aufbau von Konzepten bei Schülerinnen und Schülern unterscheidet, ... die Daten ... am besten ab“ (ebd.).

Im COACTIV-Modell „sind Wissen und Können (*knowledge*) einerseits und Werthaltungen (*value commitments*) und Überzeugungen (*beliefs*) andererseits kategorial getrennte Kompetenzfacetten“ (Baumert & Kunter, 2006, S. 496), wobei theoretisch davon ausgegangen wird, „dass die epistemologischen Überzeugungen erhebliche Bedeutung für die professionelle Wahrnehmung von Unterrichtsprozessen und das berufliche Handeln von Lehrkräften haben“ (ebd., S. 499). Die empirischen Befunde von COACTIV zeigen, dass „Lehrkräfte mit transmissiven Überzeugungen ihren Unterricht vergleichsweise gering kognitiv aktivierend und unterstützend [gestalten], während Lehrkräfte mit konstruktivistischen Überzeugungen bessere Unterrichtsqualität und entsprechend auch höhere Lernerfolge bei ihren Schülerinnen und Schülern zu verzeichnen hatten“ (Kunter & Baumert, 2011, S. 348). Transmissiv-rezeptive Überzeugungen von Mathematiklehrkräften korrelieren dabei negativ mit dem Fachwissen und dem fachdidaktischen Wissen und konstruktivistische Überzeugungen korrelieren positiv mit diesen beiden Wissensbereichen (Krauss et al., 2008b, S. 247). Ähnliche Befunde zeigen sich bei Riese und Reinhold (2010) zu Überzeugungen und Wirksamkeitserwartungen zur Physik; diese korrelieren – bis auf Überzeugungen zu einem rezeptartigen Lernen – positiv mit dem physikalischen Fachwissen und fachdidaktischen Wissen (S. 178). Diese Befunde sind konsistent zu den Ergebnissen weiterer quantitativer Studien zu berufsbezogenen Überzeugungen (vgl. u. a. Dubberke, Kunter, McElvany, Brunner & Baumert, 2008; für eine Übersicht vgl. Reusser & Pauli, 2014, S. 652f.). In der PKP-Studie werden ebenfalls epistemologische und lehr-lerntheoretische Überzeugungen³¹ erfasst (Weißeno et

³¹ Die Überzeugungen zum Lehren und Lernen enthalten die „Subfacetten ‚konstruktivistische und transmissive Überzeugungen zum Lehren und Lernen‘, ‚unterrichtsmethodische Präferenzen‘, ‚Klassenführung‘ und ‚Zielvorstellungen für den Politikunterricht‘“ (Weißeno et al., 2013, S. 195).

al., 2013, S. 194f.; Weschenfelder, 2014b, S. 149). Die empirischen Befunde zeigen, dass sich „[s]owohl bei Gym-Lehrkräften als auch bei GHR-Lehrkräften ... die hier erfassten absolutistischen Überzeugungen negativ auf Fachwissen, fachdidaktisches Wissen und unterrichtsbezogenes fachdidaktisches Wissen aus[wirken]“ (Weschenfelder, 2014a, S. 266). Ausgewählte Ergebnisse zu den lehr-lerntheoretischen Überzeugungen fassen Weißeno et al. (2013) wie folgt zusammen:

„Politiklehrer/-innen mit transmissiven Überzeugungen zum Lehren und Lernen zeigen eine Präferenz für lehrerzentrierte Lehrmethoden, bei denen die Lehrkräfte eher den aktiven Part einnehmen und der Überzeugung sind, dass Schüler/-innen am besten aus den Erklärungen und Darstellungen der Lehrerin bzw. des Lehrers lernen. Konstruktivistische Überzeugungen folgen der Vorstellung, dass Schüler/-innen politische Phänomene am besten aus den Erklärungen und Deutungen entdecken oder entwickeln. Diese beiden Facetten korrelieren deutlich negativ, was darauf hindeutet, dass Lehrer/-innen mit konstruktiven Überzeugungen eher weniger transmissive Überzeugungen zeigen, sie aber nicht entweder transmissive oder konstruktive Überzeugungen haben müssen, sondern beiden Subdimensionen zustimmen können.“ (S. 195)

Die professionellen Überzeugungen hängen dabei kaum mit dem Fachwissen und dem politikdidaktische Wissen zusammen (Weschenfelder, 2014b, S. 154f.).

Auch motivationale Orientierungen spielen für den Unterricht eine große Rolle, „damit Lehrerinnen und Lehrer ihre Kenntnisse und ihre Überzeugungen in entsprechendes Handeln überführen (vgl. Gerstenmaier und Mandl 2000; Baumert & Kunter 2006; Kunter 2011)“ (Oberle et al., 2013, S. 55). In der PKP-Studie werden von den angehenden und den bereits hauptberuflich tätigen Lehrkräften für das Fach Politik „überwiegend intrinsische Motive für die Berufswahl an[gegeben]“ (ebd., S. 60) bei gleichzeitiger positiver Ausprägung des fachlichen Interesses, was sich auch – auf geringerem Niveau – bei den Lehramtsstudierenden feststellen lässt (ebd.). Eine Quasi-Längsschnitterhebung von Gronostay und Manzel (2015) zur Studien- und Berufswahlmotivation angehender Lehrerinnen und Lehrer in Nordrhein-Westfalen bestätigt „insgesamt eine überwiegend intrinsisch geprägte Studienwahlmotivation der Befragten“ (S. 302).

Empirische Befunde zu handlungsnahen Aspekten professioneller Kompetenzen sind derzeit noch in eher geringem Umfang verfügbar. Insbesondere in der Politikdidaktik gibt es hierzu noch keine Forschungserkenntnisse. In der Mathematik konnte in der Studie von Lindmeier (2011, S. 184ff.) die aktionsbezogene Kompetenz von Mathematiklehrkräften nicht unab-

hängig vom Basiswissen (Fachwissen und fachdidaktisches Wissen) gemessen werden; das Basiswissen ist dabei eine zentrale Determinante innerhalb des Strukturmodells. Die Unterscheidung der Strukturkomponenten des Basiswissens sowie der aktionsbezogenen Kompetenz wird betont (Lindmeier et al., 2012, S. 116), da die empirischen Befunde die Aussage stützen, dass eine solche mehrdimensionale Modellierung der professionellen Kompetenz von Lehrkräften Kompetenzaspekte erfasst, die über das deklarative Wissen hinausgehen (Kniewel et al., 2015). Riese und Reinhold (2010) berücksichtigen in ihrer Modellierung und Operationalisierung ebenfalls handlungsnaher Aspekte professioneller Kompetenzen³² – hier von (angehenden) Physiklehrkräften – und kommen zu dem Schluss, „dass eine Lehrkraft im Physikunterricht zu kompetentem Handeln (im Sinne prozeduralen Wissens, soweit die verwendeten Unterrichtsvignetten hierüber Aussagen erlauben ...) imstande sein kann, sofern sie über hohes (schulbezogenes ...) Fachwissen und hohes fachdidaktisches Wissen verfügt, ein rezeptartiges oder kleinschrittiges Vorgehen zumindest kritisch hinterfragt und darüber hinaus eine konstruktivistische Sichtweise vom Lehr-Lernprozess hat“ (ebd., S. 182).

Als Determinanten der professionellen Kompetenz von angehenden Physiklehrkräften machen Riese und Reinhold (2012) unter anderem den Studiengang fest; es „konnte entsprechend ihres umfangreicheren Studiums [*sic*] durchweg ein höheres Niveau im GYM-Studiengang insbesondere bei fortgeschritteneren Studierenden beobachtet werden, wobei die Zuwächse ggf. allerdings infolge einer gewissen Stichprobenmortalität überschätzt worden sein könnten“ (S. 137). Hinsichtlich der fachbezogenen Physikkompetenz zeigen sich deutliche Geschlechtereffekte (ebd.); diese Unterschiede sind im physikdidaktischen Wissen nicht auszumachen, es sind „weder Geschlechterunterschiede noch Vorteile für HR-Studiengänge [Haupt-/Realschule] feststellbar“ (ebd.). Auch die Ergebnisse der COACTIV-Studie weisen auf eine Ausbildungsabhängigkeit der Leistungsunterschiede hin, „wie die großen Leistungsvorteile der Gymnasiallehrkräfte sowohl im Fachwissen als auch fachdidaktischen Wissen beweisen, die auch nach Kontrolle von kognitiven Eingangsvoraussetzungen bestehen bleiben“ (Kunter & Baumert, 2011, S. 347). Die differenziellen Analysen der Testleistungen der Probanden im Rahmen des PKP-Forschungsprogramms zeigen, dass die getesteten Politiklehrkräfte der Sekundarstufe II „sowohl beim fachwissenschaftlichen

³² Anders als bei Lindmeier et al. (2012) kommen zur Diagnostik handlungsnaher Aspekte professioneller Kompetenzen bei Riese und Reinhold (2010, S. 174f.) zwar auch Videovignetten zum Einsatz, hier jedoch in transkribierter Form; die verschriftlichten Unterrichtsausschnitte werden von den Probanden schriftlich in ausgewählten Aspekten analysiert und bewertet.

als auch bei den fachdidaktischen Tests besser ab[schneiden] als die Lehrer/-innen der Sek. I, wobei die Unterschiede in den fachdidaktischen Wissensfacetten geringer ausfallen“ (Weißeno et al., 2013, S. 191). Wird schließlich „für Fachwissen [statistisch kontrolliert], verschwindet der direkte Effekt der Schulform auf die fachdidaktischen Wissensfacetten“ (Weschenfelder, 2014a, S. 222). Geschlechtsspezifische Unterschiede in den Testleistungen zeigen sich auf moderatem Niveau im fachwissenschaftlichen Wissenstest (ebd., S. 225). Die Abiturnote erweist sich in der PKP-Studie nicht als Erklärungsfaktor für die Testleistungen (ebd., S. 246).

Zur Frage nach dem Effekt von Berufserfahrung auf das Professionswissen von Lehrkräften fassen Krauss et al. (2008b) zusammen:

„Die Frage, ob durch die tägliche Unterrichtspraxis Fachwissen und fachdidaktisches Wissen hinzu gewonnen wird, wird durch die COACTIV-Daten für Mathematiklehrkräfte hinreichend klar beantwortet: Statistische Analysen, die an Brunner et al. (2006b) angelehnt sind, ergeben keinerlei positiven Zusammenhang zwischen Unterrichtserfahrung (operationalisiert als Anzahl der bislang unterrichteten Jahre) und Fachwissen bzw. fachdidaktischem Wissen. Es ergeben sich im Gegenteil sogar leicht negative Korrelationen, nämlich -0.16 für das Fachwissen und -0.15 für das fachdidaktische Wissen.

Dieses Ergebnis (das also im Einklang mit der deliberate-practice-Theorie steht) legt die Vermutung nahe, dass das Wissen von Mathematiklehrkräften bezüglich dieser beiden Wissenskategorien im Wesentlichen in der Ausbildung erworben wurde.“ (S. 244f.)

Effekte von Berufserfahrung sollten sich auch in Kompetenzunterschieden zwischen Studierenden, Referendaren und Lehrkräften zeigen. Hierzu berichten Weißeno et al. (2013), dass „[d]er Unterschied, der unterrichtenden Lehrer/-innen in allen drei Wissensfacetten die höchsten Kompetenzwerte bescheinigt, ... möglicherweise auf eine Weiterentwicklung und Festigung des Wissens im Beruf hindeuten [könnte]“ (S. 192), wobei Studierende im Lehramt für die Sekundarstufe I in allen drei Wissenstests bessere Testwerte erzielen als die beiden anderen Gruppen (ebd.).

Zusammenfassend lässt sich festhalten, dass die hier referierten empirischen Befunde der Politikdidaktik sowie anderer Fächerdomänen die Annahme der zentralen Stellung des fachdidaktischen Wissens als Aspekt professioneller Handlungskompetenzen von Lehrkräften stützen. Da jedoch „fachdidaktisches Wissen ohne substantielles Fachwissen undenkbar ist“ (Baumert & Kunter, 2011b, S. 182), zeichnen sich positive Zusammenhänge zwischen den

Strukturkomponenten ab. Die Ausbildungsabhängigkeit der Leistungen zeigt sich unter anderem in der COACTIV- und in der PKP-Studie; insbesondere im Fachwissen sind die Testleistungen der Gymnasiallehrkräfte höher. Geschlechtsspezifische Effekte werden in der PKP-Studie nur bei den Gymnasiallehrkräften im Fachwissen nachgewiesen. Die Unterschiede in den Testleistungen zwischen den Studierenden und den Referendaren sowie Lehrkräften sind bei der Interpretation der Daten in der vorliegenden Studie besonders zu berücksichtigen (vgl. auch Kapitel 7.3 und Kapitel 8.1). Obwohl in dieser Arbeit mit dem politikdidaktischen Wissen und Können ein ausgewählter Kompetenzbereich professioneller Handlungskompetenzen im Fokus steht, deuten die Befunde zur Motivation und zu professionellen Überzeugungen bereits auf Anknüpfungsmöglichkeiten für Folgestudien hin (für einen Ausblick vgl. auch Kapitel 8.2).

5 Forschungsfragen

Eine *erste* Fragestellung („Ist es möglich, das politikdidaktische Wissen und Können als Bereich des Professionswissens von Politiklehrkräften in einem erweiterten Strukturmodell dazustellen?“) wurde bereits im theoretisch-explorativen Teil der vorliegenden Arbeit ergründet. Hierbei stand die Frage nach der (angenommenen) Handlungsvalidität vorliegender Modellierungen, Konzeptualisierungen und Operationalisierungen im Fokus der fächerübergreifenden sowie domänenspezifischen Analysen. Die theoretischen Erkenntnisse und vorliegenden empirischen Befunde deuten darauf hin, dass handlungsnahe Aspekte des politikdidaktischen Wissens bislang noch keiner adäquaten Erhebung zugänglich gemacht werden konnten, was durch die derzeit eher gering ausgeprägte Forschungslage zum politikdidaktischen Wissen erklärt werden kann. Da gleichzeitig aus fächergenereller (Oser et al., 2010, S. 5), aber auch aus domänenspezifischer Perspektive (Weschenfelder, 2014a, S. 79) auf die Notwendigkeit zukünftiger Forschungen zu Aspekten professionellen Wissens und Könnens von Lehrkräften hingewiesen wird, die den Handlungsbezug der zugrunde liegenden Kompetenzkonstrukte sowie eng damit verbunden die Kontextspezifität stärker betonen, wurde im weiteren Verlauf der vorliegenden Arbeit eine mehrdimensionale Modellierung des Professionswissens von Politiklehrkräften vorgeschlagen. Diese Modellierung wurde im theoretischen Anschluss an die Konzeptualisierung des Professionswissens in Nachbardomänen sowie vor dem Hintergrund genuin politikdidaktischer Überlegungen begründet. Die theoretische Konzeptualisierung der Dimensionen des Professionswissens im Rahmen der Modellierung war schließlich Ausgangspunkt für eine Ausdifferenzierung eines heuristischen mehrdimensionalen Strukturmodells des politikdidaktischen Wissens und Könnens von (angehenden) Lehrkräften. Eine Kompatibilität zu vorliegenden Konzeptualisierungen der professionellen Kompetenz von Politiklehrkräften in anderen Studien stand dabei im Fokus, um ein möglichst hohes Maß an Anschlussfähigkeit sowie Vergleichbarkeit der Ergebnisse zu ermöglichen, wobei Differenzannahmen auf Basis der aufgezeigten Forschungslücken zu berücksichtigen sind. Dies betrifft insbesondere den forschungspraktischen Zugang zur Erhebung prozeduralen politikdidaktischen Wissens, da die mehrdimensionale Modellierung erweiterte Anforderungen an eine valide Kompetenzdiagnostik stellt.

Folglich kann die *zweite* Fragestellung so formuliert werden: Kann ein Leistungstest zur Diagnostik der politikdidaktischen Kompetenz von (angehenden) Lehrkräften entwickelt

werden, der die im Modell dargestellten Facetten des politikdidaktischen Wissens und Könnens erfasst? Wegen der problematisierten Limitationen schriftlicher Erhebungsverfahren müssen zur Beantwortung dieser Frage alternative Testmöglichkeiten berücksichtigt werden (Frey & Hartig, 2013; Jurecka & Hartig, 2007), wofür in Kapitel 6.1 und Kapitel 6.2 ein computerbasierter Zugang im Format des video-cued testing beschrieben wird. Hiermit verbunden ist das Ziel, adäquate Messinstrumente (Kapitel 6.3) zu entwickeln; dies bezieht sich sowohl auf die Operationalisierung der Konstrukte zum Fachwissen (Kapitel 6.3.1.1) und zum politikdidaktischen Wissen (Kapitel 6.3.1.2) als auch auf die Kriterien zur Beurteilung der Leistungen der Probanden (Kapitel 6.3.2). Schließlich sind die Testkonstruktion selbst sowie die Expertenratings entlang der üblichen Testgütekriterien zu überprüfen (Kapitel 7.3).

Die abschließende *dritte* Fragestellung bezieht sich auf die Modellierung des politikdidaktischen Wissens und Könnens und lässt sich so zusammenfassen: Welche Zusammenhänge zeigen sich zwischen den konzeptualisierten Facetten des politikdidaktischen Wissens und wie lassen sich diese interpretieren? Es wird im Rahmen dieser Arbeit theoretisch angenommen, dass sich das fachdidaktische Wissen und das Fachwissen im Sinne eines transformativen Modells (Gess-Newsome, 1999) analytisch voneinander trennen lassen, wobei die empirischen Befunde aus der Politikdidaktik sowie anderen Fächerdomänen auf einen hohen Zusammenhang zwischen dem fachdidaktischen Wissen und dem Fachwissen hindeuten. Die in der vorliegenden Studie getrennt erhobenen Dimensionen des fachdidaktischen Wissens und des Fachwissens können zur Überprüfung dieser differenziellen Annahme miteinander in Beziehung gesetzt werden (Kapitel 7.2.6). Weiterhin wird zu überprüfen sein, welche Effekte handlungsrelevante Personenmerkmale wie bspw. *unterrichtspraktische Erfahrungen* auf das diagnostizierte politikdidaktische Wissen und Können haben und wie sich dies entlang der differenziellen Annahmen zu den einzelnen Strukturkomponenten erklären lässt (Kapitel 7.3).

6 Design der Studie

In diesem Kapitel werden zunächst die Anlage der Studie (Kapitel 6.1) und der Testablauf (Kapitel 6.2) beschrieben. Danach folgt eine Darstellung der Entwicklung der Testinstrumente (Kapitel 6.3).

6.1 Anlage der Studie

Die Studie ist als Leistungstest ausgelegt. Wie teilweise auch in den oben aufgeführten Untersuchungen wird „[d]ie standardisierte Erhebung der *Wissenskonstrukte* traditionell über Papier-Bleistift-Tests realisiert“ (Lindmeier, 2013, S. 45). Ergänzend hierzu werden handlungsnahe Aspekte des fachdidaktischen Wissens in einem computerbasierten Verfahren im Format des video-cued testing erhoben. Um einen möglichst hohen Standardisierungsgrad dieses Testbestandteils zu erreichen, werden in Anlehnung an die Vorgehensweise bei Lindmeier et al. (2012) Maßnahmen zur Kontrolle des Antwortformats sowie zur Sicherung der adäquaten testpraktischen Durchführung zum Teil technisch umgesetzt; hierauf wird in den Folgekapiteln zur Entwicklung der Testinstrumente noch näher eingegangen.

Das für die vorliegende Studie gewählte *explorative Vorgehen* kann dadurch begründet werden, dass der aus der Literaturlektüre hervorgehende domänenspezifische Theoriestand zu handlungsnahen Aspekten politikdidaktischer Kompetenzen teilweise noch gering ausgeprägt ist und dass – hiermit eng verbunden – auch aktuell vorliegende empirische Befunde ebenfalls noch nicht sehr zahlreich vorhanden sind. Für eine *explanative Untersuchung* wären gut begründete Hypothesen notwendig, deren Formulierung einen entsprechend elaborierten Stand der Forschung voraussetzt (Bortz & Döring, 2006, S. 50). Dass ein exploratives Vorgehen jedoch ebenfalls nicht ohne theoretische Vorüberlegungen erfolgt, wird betont (ebd., S. 356). Es können dabei vier Explorationsstrategien unterschieden werden, wobei es sich bei der vorliegenden Studie um eine kombinierte Strategie aus *theoriebasierter Exploration* und *explorativ quantitativer Datenanalyse* handelt (ebd., S. 358ff.). Ausgangslage einer theoriebasierten Exploration ist die Theorieanalyse mit dem nachfolgenden Arbeitsschritt, dass „Theorieansätze ... gegeneinander abgewogen und in ihrem Erklärungswert, empirischen Gehalt oder Bestätigungsgrad kontrastiert [werden]“ (ebd., S. 362), um auf dieser Grundlage schließlich – sofern möglich – die Theorien zu formalisieren und ein entspre-

chendes Modell zu bilden (ebd., S. 363). In der vorliegenden Arbeit war eine solche theoriebasierte Exploration Gegenstand des ersten Teils. Bei unterschiedlichen Ansätzen der explorativen quantitativen Datenanalyse geht es im Wesentlichen darum, vorliegende Datensätze auf Besonderheiten hin zu untersuchen oder hochdimensionale Skalen zu reduzieren, um auf dieser Basis Hypothesen für nachfolgende inferenzstatistische Verfahren entwickeln zu können. Werden beispielsweise Signifikanztests durchgeführt, so geschieht dies – unter Berücksichtigung des Umstands, dass a priori keine Hypothesen formuliert wurden – in Form *explorativer Signifikanztests* (ebd., S. 379f.). Da in dieser Arbeit keine Daten aus anderen Forschungen für die Analyse zur Verfügung stehen, werden eigene Daten erhoben, was Gegenstand dieses zweiten Teils der Untersuchung ist.

Im Sinne des Ansatzes von Lindmeier et al. (2012) lässt sich die vorliegende Arbeit zusammenfassend auch als *Machbarkeitsstudie (feasibility study)* bezeichnen, wobei der stärker erkundende Charakter der Untersuchung zu betonen ist. Der Begriff der Machbarkeitsstudie ist in den Sozialwissenschaften weniger verbreitet als beispielsweise im Bereich des Projektmanagements oder in der medizinischen Forschung. Die Aufgaben und Funktionen von Machbarkeitsstudien für die Medizin werden von Bowen et al. (2009) folgendermaßen beschrieben:

„Feasibility studies are used to determine whether an intervention is appropriate for further testing; in other words, they enable researchers to assess whether or not the ideas and findings can be shaped to be relevant and sustainable. Such research may identify not only what—if anything—in the research methods or protocols needs modification but also how changes might occur.“ (S. 453)

Zwar beschäftigt sich die vorliegende Studie entlang dieser Beschreibung nicht mit angenommenen Wirkungen von Interventionen, dennoch lässt sich diese Charakterisierung im Wesentlichen auch auf diese Arbeit beziehen.

6.2 Testablauf

An der Studie nahmen Studierende des Lehramtes Sozialwissenschaften der Universität Duisburg-Essen teil. Die Teilnahme wurde mit einer Aufwandsentschädigung von 10 Euro vergütet. Die Datenerhebung fand im Zeitraum von Ende Mai 2014 bis Mitte Dezember 2014 an der Universität Duisburg-Essen statt. Die Erhebung bestand aus einem schriftlichen

Leistungstest sowie aus einem computerbasierten Test im Kompetenzdiagnostik-Programm *vKid*³³ (Lindmeier, 2014b). Die Testdurchführung fand unter Aufsicht des Verfassers dieser Arbeit statt. Um Versuchsleitereffekte auszuschließen, folgte der Ablauf der Tests einem standardisierten Verfahren, das in einem schriftlichen Verlaufsskript vorab definiert wurde (D. Neumann, 2015, S. 4ff.)³⁴; der Testverlauf wurde vom Versuchsleiter protokolliert, um mögliche Zwischenfragen, auftretende technische Probleme oder andere Unregelmäßigkeiten dokumentieren zu können (Bortz & Döring, 2006, S. 84). Während des Tests war der Versuchsleiter nicht anwesend, wobei er von den Probanden bei Fragen in einem anderen Raum kontaktiert werden konnte.

Der Test ist in unterschiedliche Blöcke unterteilt, die an die Logik der entwickelten Fragebatterien anschließen. Neben den Fragen zum Fachwissen sowie zum fachdidaktischen Wissen werden Hintergrundinformationen bei den Probanden erfragt, wie z. B. *Alter, Geschlecht, Abiturnote, Angaben zum Studienverlauf und Unterrichtserfahrungen*. ‚Eisbrecher-Fragen‘ zum Beginn der Frageblöcke sollen einen Motivationsabfall verhindern (ebd., S. 244); um Ermüdungseffekte (ebd., S. 193) zu minimieren, sind Fragen mit visuell anregenden Bestandteilen in den Fragebatterien verteilt aufgenommen. Der Frageteil zum Fachwissen wird vor den Fragen zum fachdidaktischen Wissen bearbeitet. Nach den Fragen zum deklarativen politikdidaktischen Wissen folgt abschließend der Aufgabenteil zum prozeduralen politikdidaktischen Wissen im Format des video-cued testing. Die Probanden arbeiten hierfür in einer computergestützten Testumgebung. Die Ausgabe des Tonsignals und die Aufzeichnung der mündlichen Rückmeldung erfolgen kombiniert über ein *Headset*; das Video wird in einem vergrößerbaren Videofenster ausgegeben. Auf einem separaten Notizklemmbrett können sich die Testpersonen optional ergänzende Aufzeichnungen machen, die ihnen bei der Bearbeitung der Aufgaben hilfreich sein können; diese Aktivitäten werden mit einer kompakten Videokamera aufgezeichnet. Die Teilnehmenden selbst werden nicht video-grafiert (Abbildung 16).

³³ Das Programm *vKid* (Videobasierte Kompetenz- und Instruktionsdiagnostik) wurde unter der Leitung von Prof. Dr. Anke Lindmeier an der Technischen Universität München entwickelt (TUM, 2015).

³⁴ Die Unterlagen zur Organisation der Studie inklusive Dokumentation der verwendeten Erhebungsinstrumente, Ratingmanuale und Datenaufbereitung (D. Neumann, 2015) können beim Verfasser der vorliegenden Arbeit angefordert werden.



Abbildung 16: Aufbau der Testumgebung mit Laptop, akustischem Ausgabe-/Aufnahmegerät (Headset) und Notizklemmbrett mit Videoperipherie

Wegen des Laborcharakters des Tests wurde auf möglichst gleichbleibende Testbedingungen geachtet (ebd., S. 84). Alle Testdurchläufe fanden im gleichen Raum (*Didaktiklabor*) an denselben technischen Haupt- und Peripheriegeräten statt. Aus forschungsökonomischen Gründen konnten die Tests nicht zur selben Uhrzeit stattfinden, jedoch wurden Testdurchläufe stets innerhalb der für die Probanden gewohnten Wochenstudienzeiten terminiert (wochentags im Zeitraum von 9.00 Uhr bis spätestens 18.00 Uhr).

6.3 Entwicklung der Testinstrumente

Bei der Entwicklung der Testinstrumente ist zu berücksichtigen, dass es sich bei Kompetenzen um Personenmerkmale handelt, die nicht direkt beobachtbar sind und daher nur indirekt

gemessen werden können (Bühner & Ziegler, 2009, S. 15). Um die Ausprägungen dieser latenten Personenmerkmale erfassen zu können, ist der Schritt der Operationalisierung notwendig. Durch diesen (in der vorliegenden Arbeit in Form der Konstruktion von Fragen [Test-Items] für den Leistungstest) „lässt sich die Ausprägung jeder Person in dem Merkmal als eine Zahl ausdrücken“ (ebd., S. 16). Bei der Entwicklung von Testinstrumenten müssen zunächst die Aufgabeninhalte festgelegt werden (Hartig & Jude, 2007, S. 26), was die Auswahl der erwünschten Antwortformate umfasst. Neben geschlossenen Antwortformaten (Auswahl einer oder mehrerer Lösungen) und offenen Antwortformaten (häufig freie Textfelder) (ebd., S. 29f.) können halb offene Antworten zur Anwendung kommen. Entsprechende Items „überlassen die Antwortformulierung dem Untersuchungsteilnehmer; die gestellte Aufgabe sollte jedoch im Unterschied zu einem offenen Item so präzise sein, dass nur eine Antwort richtig ist“ (Bortz & Döring, 2006, S. 213). Die Vor- und Nachteile der geschlossenen und offenen Antwortformate stehen sich im Prinzip gegenüber: Während offene Items aufwendig zu konstruieren sind, da ‚passende‘ Distraktoren gefunden werden müssen, sind sie gleichzeitig relativ einfach und objektiv auszuwerten (Hartig & Jude, 2007, S. 30). Demgegenüber sind offene Items relativ leicht zu konstruieren, dafür aber zeitintensiver in der Auswertung (ebd.). Bortz und Döring (2006) diskutieren außerdem qualitative Unterschiede in den Fähigkeiten, die beide Antwortformate erfordern. So sei die Qualität der gezeigten Fähigkeit zur Beantwortung bei Multiple-Choice-Fragen niedriger zu bewerten als bei freien Antworten (S. 215).

In der vorliegenden Studie kommt jedes der aufgeführten Antwortformate in unterschiedlichen Gewichtungen vor. Die Entscheidung für oder gegen ein bestimmtes Format wird dabei maßgeblich von zwei Kriterien beeinflusst: erstens von der inhaltlichen Passgenauigkeit in Bezug auf das erwünschte Antwortverhalten und zweitens von testökonomischen Überlegungen. Hinsichtlich des ersten Kriteriums wird davon ausgegangen, dass sich deklaratives Wissen über politikdidaktische Fakten und Sachverhalte in erster Linie als reproduktives Wissen von den Testpersonen abrufen lässt und daher in Form von geschlossenen Items im Test abgefragt werden kann. Halb offene Items werden unter anderem dann dargeboten, wenn die Teilnehmenden für die Beantwortung eher unterrichtsnahes Wissen aktivieren oder grundlegende problemlösende Strategien anwenden müssen. Die Test-Items, in denen die Probanden auf einen Videostimulus mit einer mündlichen Rückmeldung reagieren müssen, erfordern in der vorliegenden Studie schließlich Frageformulierungen, die offene Antworten

ermöglichen. Bei der Konstruktion von geschlossenen Items ist die Anzahl der Distraktoren ein kritisches Moment, da richtige Antworten auch erraten werden könnten (Bortz & Döring, 2006, S. 216f.). Während lediglich eine Antwortalternative (richtig/falsch) aufgrund der Trefferwahrscheinlichkeit von 50 % als zu wenig erachtet wird (ebd., S. 216), kann eine Anzahl von drei Distraktoren als ausreichend angesehen werden, um die Ratewahrscheinlichkeit zu vermindern.

Ausgangspunkt für die Entwicklung der Testinstrumente der Studie ist das mehrdimensionale Strukturmodell politikdidaktischen Wissens und Könnens der vorliegenden Arbeit (Abbildung 14). Durch die inhaltsbezogene Konzeptualisierung zum Themenbereich *Parteienverbote in der Parteiendemokratie am Beispiel der NPD-Verbotsverfahren* wird in Abbildung 17 die Verortung des Schwerpunktes der vorliegenden Studie nochmals grafisch im zugrunde liegenden Strukturmodell dargestellt.

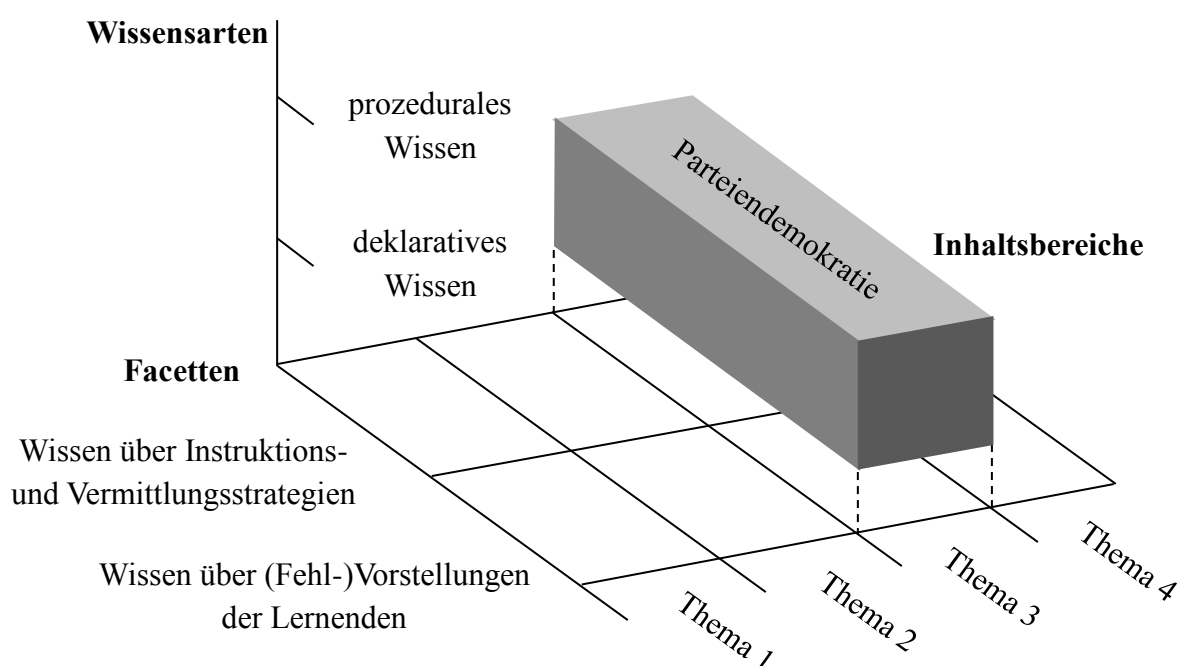


Abbildung 17: Heuristisches mehrdimensionales Strukturmodell zum politikdidaktischen Wissen und Können mit Spezifizierung der Inhaltsdimension in der vorliegenden Studie (eigene Darstellung in Anlehnung an Schmelzing, 2010; Tepner et al., 2012, Weißeno et al., 2013 sowie ergänzend Faissner et al., o. J., S. 14)

6.3.1 Operationalisierung der Konstrukte

Es können unterschiedliche Möglichkeiten zur Konstruktion von Testinhalten zusammengefasst werden (für eine Übersicht vgl. Hartig & Jude, 2007, S. 26ff.), wobei zwei Strategien als besonders zielführend für kompetenzdiagnostische Verfahren erachtet werden: die *deduktive Konstruktion* und das *Kriteriumssampling* (ebd.). Bei der deduktiven Konstruktion werden „[a]uf Basis der Beschreibung des theoretischen Konstrukts ... mögliche beobachtbare Indikatoren abgeleitet, in denen sich interindividuelle Unterschiede im Konstrukt möglichst stark niederschlagen sollten“ (ebd., S. 27). Beim Kriteriumssampling „wird aus einem interessierenden Verhaltensbereich, in dem erfolgreiches Handeln vorhergesagt werden soll, eine möglichst repräsentative Stichprobe von Verhaltensweisen ausgewählt“ (ebd., S. 28) und „dann in Testaufgaben überführt“ (ebd.). Ein solches Kriteriumssampling wird als besonders zielführende Strategie zur Itementwicklung für ein kompetenzdiagnostisches Verfahren angesehen, „wenn man Kompetenzen als kontextspezifische, durch einen bestimmten Situationsbereich definierte Konstrukte betrachtet“ (ebd.).

In dieser Arbeit wird in Anlehnung an das Vorgehen bei Lindmeier (2011, S. 126f.) bei der Entwicklung von Test-Items eine kombinierte Strategie aus deduktiver Konstruktion und Kriteriumssampling verfolgt. Bei einer deduktiven Vorgehensweise ist Ausgangslage für die Entwicklung von Testaufgaben das Modell politikdidaktischen Wissens und Könnens von Politiklehrkräften, dessen Konzeptualisierungen der Facetten sowie dessen Strukturzusammenhang zu einem übergeordneten Kompetenzmodell theoretisch begründet sind. Das Modell beschreibt auf Basis beruflicher Anforderungssituationen Personenmerkmale, die notwendig sind, um diese Anforderungen professionell bewältigen zu können. Insofern lassen sich „[a]us diesem Modell ... Aufgaben ableiten, die von kompetenten Personen mit höherer Wahrscheinlichkeit gelöst werden können als von einer weniger kompetenten Person“ (Hartig & Jude, 2007, S. 29). Da zur Diagnostik handlungsnaher Aspekte gleichzeitig authentische Anforderungssituationen als Grundlage für die Aufgabenformulierung herangezogen werden, deren Auswahl zudem entlang von ex ante formulierten Kriterien zur Beschreibung hoher Qualität des Unterrichtens erfolgt, lassen sich bei diesem Vorgehen Merkmale eines Kriteriumssamplings erkennen (Lindmeier, 2011, S. 126f.), da „[d]ie Begründung für einen Aufgabeninhalt ... darauf [basiert], dass die Aufgabe eine inhaltliche Entsprechung zu realen Situationen im interessierenden Bereich darstellt“ (Hartig & Jude, 2007, S. 29).

Die Entwicklung von Test-Items für die vorliegende Studie orientiert sich am Vorgehen des COACTIV-Forschungsprogramms, bei dem Items „mithilfe von Literaturrecherchen, etwa durch Inspektion und Weiterentwicklung von Aufgaben aus den internationalen Vergleichstests“ (Krauss et al., 2008b, S. 239) sowie „aus der ‚mathematikdidaktischen mündlichen Tradierung‘“ (ebd.) zusammengetragen wurden. Für die Politikdidaktik sind jedoch eigene domänenspezifische Überlegungen notwendig, da bei der Orientierung an Testaufgaben aus Nachbardomänen zu bedenken ist, dass „[d]ie Übernahme solcher Konstrukte ... gleichwohl die politikdidaktische Professionsforschung bildungswissenschaftlich anschlussfähig [macht], [sie] freilich ... nolens volens auch sehr viele Ergebnisse [generiert], die bislang vor allem domänengenerell sind (hier in weitgehender Anlehnung an MT21/TEDS-M Blömeke/Kaiser/Lehmann 2008; Oberle u.a. 2012)“ (Szukala, 2013, S. 33f.).

6.3.1.1 Entwicklung der Test-Items zum Fachwissen

Da das Fachwissen entlang der theoretischen Annahmen des transformativen Modells eine analytisch vom fachdidaktischen Wissen trennbare Kompetenzfacette darstellt, erfolgt dessen Erhebung durch separate Test-Items. Hierzu wurden zunächst die vorliegenden Planungsunterlagen der Lehrkräfte mit Fokus auf den in den videografierten Unterrichtsstunden verwendeten Unterrichtsmaterialien untersucht. Die Unterrichtsstunden wurden im Rahmen der Studie *Argumentative Lehr-Lernprozesse* (Gronostay, 2015) aufgezeichnet. Alle Stunden behandeln das Thema *Parteienverbote in einer Parteiendemokratie*. Es wurde hierbei das Schülerkompetenzmodell von Detjen et al. (2012) in der Kompetenzdimension des Fachwissens (Weißeno et al., 2010) herangezogen, um zu ermitteln, welches konzeptuelle Wissen mithilfe welcher konstituierenden Fachbegriffe im Unterrichtsverlauf bei den Schülerinnen und Schülern in dieser Jahrgangsstufe induziert werden sollte (D. Neumann, 2015, S. 62ff.).

Auf Grundlage dieser Analyse wurden in mehreren Expertensitzungen zur Itemkonstruktion entsprechende Testaufgaben zum themenbezogenen Fachwissen entwickelt, wobei hinsichtlich des Niveaus davon ausgegangen wurde, dass, „[u]m herausfordernden Situationen im Unterricht fachlich gewachsen zu sein und Argumentationsweisen und herzustellende Zusammenhänge an typische Wissensbildungsprozesse des Fachs anschließen zu können, ... Lehrkräfte den Unterrichtsstoff auf einem höheren Niveau als dem der unterrichtlichen Bearbeitung zu durchdringen [haben] (vgl. Krauss et al., 2011, S. 142f.)“ (Weschenfelder,

2014a, S. 166). Die Test-Items sind daher auf einem Niveau konzeptualisiert, das als „[t]ieferes Verständnis der Fachinhalte des Curriculums der Sekundarstufe“ (Krauss et al., 2008b, S. 237) bezeichnet werden kann (vgl. auch Weschenfelder, 2014a, S. 166). Zudem wurde bei der Entwicklung der Test-Items berücksichtigt, dass trotz der Themenspezifität das Fachwissen innerhalb einer gewissen Breite abgebildet wurde, was auch insofern wichtig ist, als die Vernetzungsansätze beim Aufbau des konzeptuellen Wissens durchgängig betont werden (Manzel & Gronostay, 2013, S. 201ff.; Weißeno et al., 2010, S. 21f.). Alle 40 Test-Items zum Fachwissen sind in einem geschlossenen Antwortformat (Multiple-Choice-Format mit drei Distraktoren) abgebildet und werden aus testpraktischen Gründen in einem separaten ‚traditionellen‘ Papier-Bleistift-Format bearbeitet (D. Neumann, 2015, S. 10ff.) (Abbildung 18).³⁵

Was ist kein Merkmal von Rechtsstaatlichkeit?

- ☐ Gesetzmäßigkeit der Verwaltung
- ☐ Ausübung staatlicher Macht auf Grundlage der Verfassung
- ☒ Entscheidung über Gesetze durch die Bürger/-innen
- ☐ Vorbehalt des Gesetzes

Abbildung 18: Beispiel-Item (geschlossenes Antwortformat) zur Erhebung des Fachwissens

6.3.1.2 Entwicklung der Test-Items zum politikdidaktischen Wissen

Die bereits angesprochenen Problemstellungen bei der Entwicklung von Test-Items werden beim politikdidaktischen Wissen besonders deutlich. So zeigt eine Zusammenschau über politikwissenschaftliche und politikdidaktische Vertiefungsteile des Studiums an unterschiedlichen Hochschul- und Universitätsstandorten in Baden-Württemberg, Nordrhein-

³⁵ Hintergrund hierfür ist der Umstand, dass für die Befragungen lediglich ein Computer mit entsprechenden Peripheriegeräten zur Verfügung steht. Um bei direkt aufeinanderfolgenden Testteilnahmen einen gewissen zeitlichen Spielraum zu haben, können nachfolgende Probanden bereits mit diesem fachwissenschaftlichen Fragebogen beginnen, selbst wenn eine Testperson noch den computerbasierten Test bearbeitet; diese Befragungsteile finden in jeweils unterschiedlichen Räumen statt.

Westfalen und Bayern von Weschenfelder (2014a), „dass die Inhalte des politikdidaktischen Studiums im Gegensatz zum politikwissenschaftlichen an den Hochschulstandorten gering strukturiert sind“ (S. 165). Trotz bildungsadministrativer Vorgaben zu ländergemeinsamen inhaltlichen Vorgaben der Lehrerbildung können offenbar keine standardbasierten Studienbestandteile ausgemacht werden, was auch mit dem hohen Allgemeinheitsgrad inhaltsbezogener Standardformulierungen zusammenhängen dürfte (Massing, 2013, S. 275f.). Eine ausschließlich deduktive Vorgehensweise bei der Entwicklung von Test-Items zum politikdidaktischen Wissen – ausgehend von den lediglich basalen Formulierungen in den ausbildungsbezogenen Standards und/oder konkreten, jedoch bereits zwischen unterschiedlichen Hochschulstandorten variierenden fachdidaktischen Studienelementen – scheint im Rahmen dieser Studie nicht zielführend zu sein. Stattdessen wird auch hier eine kombinierte Vorgehensweise aus normativ-deduktivem Zugang und Kriteriumssampling gewählt.

Die Konzeptualisierung erfolgt hierbei in Anlehnung an das Vorgehen im PKP-Forschungsprogramm durch Aufgliederung in unterschiedliche Aspekte der beiden Facetten, wobei für das Wissen über Instruktions- und Vermittlungsstrategien der Fokus „auf normative politikdidaktische Konzeptionen und Diskurse und auf unterrichtsbezogene Aspekte [gerichtet ist]“ (Weschenfelder, 2014a, S. 169). Auf Grundlage des thematischen Schwerpunktes (*Parteienverbote in einer Parteiendemokratie*; vgl. auch Kapitel 6.3) werden Fragen zu Mikro-, Makro- und Mesomethoden konstruiert, die sich entlang des fachdidaktischen *Kontroversitätsprinzips* formulieren lassen (vgl. auch Kapitel 2.2.2). Hierbei liegt in Anlehnung an Theorien zur kognitiven Dissonanz (Festinger, 2012) der Fokus insbesondere auf methodischen Zugängen, die als „Dissonance Strategies“ (Clement, 2008, S. 421) verstanden werden. Diese sollen die Lernenden kognitiv anregen und dabei unterstützen, alternative Sichtweisen wahrzunehmen und im Unterricht aufgeworfene Fragen und Problemlagen multiperspektivisch zu analysieren. Die Test-Items zum Wissen über (Fehl-)Vorstellungen liegen im Schwerpunkt auf Formulierungen, die „sich auf den Konzeptaufbau, auf Fördermöglichkeiten im Politikunterricht und auf die Diagnose typischer Fehlkonzepte von Schüler/-innen im Fach [beziehen]“ (ebd.). Hierzu finden die Überlegungen von Weißeno et al. (2010) über mögliche Fehlkonzepte zu den hier inhaltlich relevanten Fachkonzepten Berücksichtigung; ergänzend werden exemplarische (Fehl-)Vorstellungen aus den videografierten Unterrichtsstunden abgeleitet.

In mehreren Expertensitzungen zur Itemkonstruktion wurden unter Berücksichtigung der Befunde einer pilotierenden Befragung³⁶ insgesamt 20 Test-Items zum *deklarativen politikdidaktischen Wissen* (weiter-)entwickelt (Multiple-Choice-Fragen mit jeweils drei Distraktoren im Computerprogramm vKid); sechs Aufgaben sind halb offen formuliert (vgl. auch Abbildung 19, Abbildung 20 und Abbildung 21 für Beispiel-Items zum deklarativen politikdidaktischen Wissen).

The screenshot shows the vKIDPlayer 1.5 software interface. At the top, there is a menu bar with 'Datei' and 'Administration ?'. Below the menu bar, the title 'Fachdidaktik' is displayed. The main content area contains the question: 'Welche Aufgabe haben didaktische Prinzipien in der politischen Bildung unter anderem? Die Auswahl und Begründung von ...'. Below the question, there are four radio button options: 'Kategorien', 'Kompetenzen', 'Methoden', and 'Konzepten'. A 'Weiter' button is located at the bottom right of the question area. At the very bottom of the window, a status bar displays: 'Frage 17 von 48', 'Maximale Bearbeitungszeit: 1:30', 'Aktuelle Bearbeitungszeit: 0:08', and 'Empfohlene Bearbeitungszeit: 0:45'.

Abbildung 19: Beispiel-Item (geschlossenes Antwortformat) zur Erhebung des deklarativen politikdidaktischen Wissens über Instruktions- und Vermittlungsstrategien im Computerprogramm vKid

³⁶ Die Befragung wurde im Wintersemester 2012/13 an der Universität Duisburg-Essen im Rahmen zweier politikdidaktischer Seminare im Hauptstudium des Studienganges *Lehramt Sozialwissenschaften* (Gelegenheitsstichprobe, $N = 50$) durchgeführt. Die Befragung einer solchen nichtprobabilistischen Stichprobe erlaubt keine Rückschlüsse auf die Grundgesamtheit (Generalisierung). Die reliabilitätsstatistischen Auswertungen der Befragungen (hier insbesondere die Befunde zu Itemschwierigkeit und Item-Trennschärfe) sollen jedoch bei der weiteren Item-Entwicklung Berücksichtigung finden.

The screenshot shows a software window titled "vKIDPlayer 1.5". Inside, there is a header bar with "Datei" on the left and "Administration ?" on the right. The main content area contains the text: "Nennen Sie in Stichpunkten den Dreischritt zur Karikaturenanalyse (politikdidaktische Hermeneutik)." Below this text is a large, empty rectangular box for the user's answer. At the bottom left of the answer box, there is a small icon of a pencil and the text "Noch verbleibende Zeiche... 100". At the bottom right of the main area is a button labeled "Weiter". The footer of the window displays the following information: "Frage 25 von 48", "Maximale Bearbeitungszeit: 2:00", "Aktuelle Bearbeitungszeit: 0:09", and "Empfohlene Bearbeitungszeit: 1:00".

Abbildung 20: Beispiel-Item (halb offenes Antwortformat) zur Erhebung des deklarativen politikdidaktischen Wissens über Instruktions- und Vermittlungsstrategien im Computerprogramm vKid

The screenshot shows a software window titled "vKIDPlayer 1.5". Inside, there is a text area with the following content:

Sie diskutieren die Frage, wann eine Partei verboten werden kann.

Eine Schülerin sagt: "Wenn eine Partei demokratische Prinzipien nicht anerkennt, ist sie verfassungsfeindlich. Dann kann sie verboten werden."

Welche Rückmeldung geben Sie der Schülerin?

Below the text are four radio button options:

- ☐ "Das ist eine sehr gute Antwort."
- ☐ "Das gilt aber nur für Parteien, die in Landtagen oder im Bundestag vertreten sind."
- ☐ "Es geht nicht darum, ob sie 'verfassungsfeindlich', sondern 'verfassungswidrig' ist, also ob sie aktiv die demokratische Ordnung beseitigen will."
- ☐ "Parteienverbote richten sich danach, ob die Parteispitze an strafbaren Handlungen beteiligt war."

At the bottom right is a button labeled "Weiter". At the bottom left, it says "Frage 31 von 48". At the bottom center, it shows "Maximale Bearbeitungszeit: 2:00" and "Aktuelle Bearbeitungszeit: 0:15". At the bottom right, it shows "Empfohlene Bearbeitungszeit: 1:00".

Abbildung 21: Beispiel-Item (geschlossenes Antwortformat) zur Erhebung des deklarativen politikdidaktischen Wissens über (Fehl-)Vorstellungen der Schüler/-innen im Computerprogramm vKid

Für die Erfassung des *prozeduralen politikdidaktischen Wissens* wird davon ausgegangen, dass sich handlungsnahe Aspekte dieser Facette im Unterricht in der Interaktion zwischen Lehrenden und Lernenden zeigen, zum Beispiel dadurch, dass es einer Lehrkraft gelingt, auf eine falsche Äußerung einer Schülerin bzw. eines Schülers mit einer erklärenden Rückmeldung zu reagieren oder bei Unterrichtsmethoden wie einer Pro-Kontra-Debatte mit einem mündlichen Impuls dergestalt zu unterstützen, dass der Gesprächsfluss angeregt wird. Die Operationalisierung dieser Facette erfolgt dabei in Anlehnung an die Vorgehensweise bei Lindmeier (2011), die als Rückmeldungsformate „*Explanation*“ (S. 130) (*Erklärung*) und „*Stimulus*“ (ebd., S. 131) (in der vorliegenden Arbeit übersetzt mit *unterstützender Impuls*) unterscheidet. Hierbei kennzeichnet eine erklärende Rückmeldung eine Aufgabestellung, „[that] requires a spontaneous diagnosis of student errors in order to explain the topic spontaneously in a student-centered way“ (ebd., S. 130). Im Aufgabenformat zum unterstützen-

den Impuls sollen die Probanden den Schülerinnen und Schülern „spontaneous assistance ... in the form of a helpful stimulus“ (ebd., S. 131) geben. Dass beiden unterrichtlichen Strategien zur Rückmeldung an die Schüler/-innen dabei kein ‚mechanistisches‘ Wirkverständnis zugrunde liegt, drückt Leinhardt (2010) für das Format der *instructional explanations* so aus:

„In an effective instructional explanation, the query or problem itself must be carefully unpacked and examined, not just stated – this means identifying features in the query that are problematic and summoning effective and important examples. The knowledge system that supports the framing of a significant query is a combination of deep disciplinary knowledge and solid pedagogical understanding“ (S. 4)

In der vorliegenden Arbeit wird der Auffassung gefolgt, dass sich eine erfolgreiche Erklärung präzise auf das (Fehl-)Verständnis und die Lernvoraussetzungen der Lernenden beziehen muss, um sie kognitiv zu aktivieren (Leuders & Holzäpfel, 2011, S. 3). Aus dieser Perspektive wird auch der ko-konstruktive Charakter einer erklärenden Rückmeldung betont, der Lerner „[is] positioned as a partner in constructing understanding (being explained *with*)“ (van de Sande & Greeno, 2010, S. 69). Ein durchgehender Bezug auf das (Fehl-)Verständnis der Lernenden ist dabei noch offensichtlicher beim Format eines unterstützenden Impulses notwendig, da die Lehrperson hierbei die richtige Antwort oder bei komplexeren Fragestellungen eine adäquate Lösungsstrategie nicht vollständig verraten möchte, sondern die Schüler/-innen nur soweit unterstützen möchte, dass sie selbstständig auf einen Lösungsansatz kommen; Abbildung 22 zeigt beispielhaft den Aufbau eines Items zum Aufgabentyp *unterstützender Impuls*.



Abbildung 22: Beispiel-Item zur Erhebung des prozeduralen politikdidaktischen Wissens (Aufgabentyp *unterstützender Impuls*) im Computerprogramm vKid

Anmerkung: Aus Datenschutzgründen ist das Vorschaufenster für das Video verfremdet.

Der Videostimulus zur Erhebung des prozeduralen politikdidaktischen Wissens enthält in beiden Aufgabentypen *Erklärung* und *unterstützender Impuls* (jeweils 5 Items) eine konkrete unterrichtliche Anforderungssituation in Form der Darstellung von Verständnisschwierigkeiten der Schüler/-innen, inhalts- oder methodenbezogenen Problemen bei der Durchführung unterrichtlicher Arbeitsaufträge oder Äußerungen der Schüler/-innen, die Fehler enthalten und daher eine Hilfestellung der Lehrkraft erfordern. Die Videovignetten wurden auf Grundlage der vorliegenden Unterrichtsaufnahmen entwickelt und zeigen daher ein reales Unterrichtsgeschehen. Die Auswahl der Vignetten erfolgte in einer Expertensitzung unter der Prämisse, dass sie Problemstellungen entlang der Kriterien guten Unterrichtens enthalten (Rehm & Bölsterli, 2014, S. 219f.). Es kann zwar davon ausgegangen werden, dass solche Vignetten – schon wegen des zeitlichen Umfangs – grundsätzlich komplexitätsreduziert sind, dass sie

gleichzeitig aber durch ihre Authentizität keine für den jeweiligen Unterrichtsausschnitt relevanten Aspekte auslassen.³⁷ Die so entwickelten Vignetten wurden mit Studierenden sowie Lehrkräften in einer Fortbildung nochmals hinsichtlich der für die Testdurchführung relevanten Aspekte (z. B. akustische Verständlichkeit) analysiert, um problematische Aufzeichnungen aussortieren zu können. Daneben wurde erprobt, welche Bearbeitungszeit maximal je Item anzusetzen ist. Lindmeier (2011, S. 131) legt eine maximale Bearbeitungszeit von 2 Minuten und 30 Sekunden (ohne Videovignette) für die Beantwortung der Items im Aufgabentyp *Erklärung* fest; die maximale Bearbeitungszeit für den Aufgabentyp *Stimulus* beträgt eine Minute ohne Videovignette (ebd., S. 132). Aufgrund der geringen zeitlichen Spannweite der eingesetzten Mitschnitte in der vorliegenden Studie – die Vignetten sind 11 bis 39 Sekunden lang und haben eine durchschnittliche Länge von 25 Sekunden – zeichneten sich hierzu in präpilotierenden Versuchen mit Testpersonen nur geringe zeitliche Abweichungen in einer Antwortzeit um 3 Minuten ab. Diese Zeit wurde daher als maximale Bearbeitungszeit inklusive Videovignettenimpuls je Test-Item festgelegt.

Als übergeordnete Aufgabenstellung erhalten die Probanden die Anweisung, in beiden Aufgabentypen eine spontane Rückmeldung zu geben, die fachlich korrekt, präzise und konkret angepasst an die jeweiligen Schwierigkeiten ist und die darüber hinaus insbesondere im Aufgabentyp *Erklärung* kurz und prägnant und im Aufgabentyp *unterstützender Impuls* dergestalt erfolgen soll, dass die richtige Lösung nicht komplett verraten wird (ebd., S. 131f.). Die Antwort soll in natürlicher Sprache gegeben werden (offenes Antwortformat) und die Schüler/-innen direkt adressieren, also nicht in einer metakognitiven Formulierung (,Ich würde dem Schüler sagen, dass ...‘), sondern in unmittelbarer Ansprache (,Das ist ein guter Einwand, aber du solltest bedenken ...‘), als wenn sich die Testperson in der Unterrichtssituation befinden würden. Hierdurch wird eine höhere Standardisierung hinsichtlich der *Unmittelbarkeit* der Anforderungssituation angestrebt (Lindmeier, 2013, S. 55). Die Teilnehmenden erhalten vorab Hintergrundinformationen zur Klasse (Jahrgangsstufe, Schulform, zu erwartender Lernstand im Inhaltsbereich, bisher bearbeitete Unterrichtsmaterialien). Im jeweiligen Item wird darüber hinaus eine kurze schriftliche Beschreibung der Anforderungssituation gegeben, die die Probanden lesen, bevor sie sich das Video anschauen. Hierbei ist zu berücksichtigen, „dass Kontexte selbst zu einem schwierigkeitsgenerierenden Merkmal werden können“ (Mayer & Wellnitz, 2014, S. 27), weshalb diese Beschreibung ausschließlich ergän-

³⁷ Zu weiteren Problem bei der Auswahl von Videovignetten vgl. Brovelli et al. (2013, S. 310f.).

zende Hinweise enthält, die zum besseren Verständnis der Situation notwendig sind (Riese & Reinhold, 2014, S. 262). Im Test-Item in Abbildung 22 wird beispielsweise der Hinweis gegeben, dass sich die Verständnisschwierigkeit des Schülers in der Videovignette auf die Kernaussage einer zuvor analysierten Karikatur bezieht.

Die Bearbeitungszeit der Test-Items ist begrenzt. Sobald die Testpersonen eine neue Frage öffnen, beginnt die Bearbeitungszeit abzulaufen. Im unteren Teil des Fensters des Computerprogramms sind Angaben zur maximalen und empfohlenen Bearbeitungszeit sowie zum aktuellen Bearbeitungsstand zu sehen. Wenn die aktuelle Bearbeitungszeit die empfohlene überschreitet, wird erstere rot. Hierdurch erhalten die Probanden einen visuellen Impuls, sich zu beeilen. Nach Ablauf der maximalen Bearbeitungszeit wird die Frage geschlossen, es ist dann keine Antwort für das jeweilige Item mehr möglich. Durch den Aufbau von Zeitdruck wird ein höherer Standardisierungsgrad hinsichtlich der *Spontaneität*, die für eine Reaktion auf die gezeigte Anforderungssituation notwendig ist, ermöglicht (Lindmeier, 2013, S. 54).

6.3.2 Beurteilungskategorien und Ratingverfahren

Von den insgesamt 70 Fragen (ohne Fragen zu Hintergrundvariablen), die die Probanden im Rahmen der vorliegenden Studie beantwortet haben, ist ein Großteil geschlossen (54 Items) oder halb offen (6 Items) formuliert. Die 10 Fragen mit videovignettenbasiertem Stimulus haben eine offene Aufgabenformulierung mit konkreten Anweisungen zur erwarteten Handlung (ebd., S. 53). Wegen der Auswertungsobjektivität von Items mit Antwortvorgaben (Bortz & Döring, 2006, S. 215) wird davon ausgegangen, dass die Beurteilung durch geschulte Experten erfolgen kann. Grundsätzlich wird dies auch für halb offene Antworten angenommen, sofern die Aufgabenstellung präzise formuliert ist (ebd., S. 213). Um jedoch eine eindeutige intersubjektive Auswertung zu sichern, liegt für die Beurteilung der halb offenen Antworten ergänzend ein Ratingmanual vor (D. Neumann, 2015, S. 19ff.). Die Antworten in der Skala zum deklarativen politikdidaktischen Wissen wurden von zwei geschulten Ratern beurteilt. Nach der unabhängigen Erstbeurteilung wurden Beurteilungsabweichungen diskutiert (*consensus expert scoring*); auf dieser Grundlage wurden die halb offenen Fragen schließlich in einem dichotomen Format kodiert (richtig/falsch).

Während bei geschlossenen Antworten die Beurteilung in der Regel also ohne Schwierigkeiten möglich ist, sind Beurteilungskriterien erforderlich, um die Qualität der Antworten auf

offene Fragen beurteilen zu können: „If test items are provided without information on rating and scoring methods, their intentions can only be evaluated to a certain extent“ (Lindmeier, 2011, S. 140). Die Beurteilung wäre intransparent und würde deutlich von der individuellen Sichtweise des Raters abhängen, wodurch das Ergebnis nicht vergleichbar wäre. Insofern „bedeutet gerade das Arbeiten mit einem Kategoriensystem einen entscheidenden Punkt der Vergleichbarkeit der Ergebnisse, der Abschätzung der Reliabilität der Analyse“ (Mayring, 2010, S. 50). Auch die Validität der Urteile wird durch die Beurteilung von mehreren Ratern erhöht (Bortz & Döring, 2006, S. 185).

6.3.2.1 Entwicklung der Ratingkategorien und des Ratingmanuals

Die grundlegende Bildung der Beurteilungskategorien orientiert sich an dem von Mayring (2010) in diesen drei Schritten zusammengefassten Verfahren:

- „1. Definition der Kategorien
Es wird genau definiert, welche Textbestandteile unter eine Kategorie fallen.
2. Ankerbeispiele
Es werden konkrete Textstellen angeführt, die unter eine Kategorie fallen und als Beispiele für diese Kategorie gelten sollen.
3. Kodierregeln
Es werden dort, wo Abgrenzungsprobleme zwischen Kategorien bestehen, Regeln formuliert, um eindeutige Zuordnungen zu ermöglichen.“ (S. 92)

Für die vorliegende Arbeit werden die Kategorien in ihren wesentlichen Aspekten *deduktiv* gebildet. Im Gegensatz zu einer *induktiven Kategoriendefinition*, die „Kategorien direkt aus dem Material in einem Verallgemeinerungsprozess ab[leitet], ohne sich auf vorab formulierte Theorienkonzepte zu beziehen“ (ebd., S. 83), wird bei einer *deduktiven Kategoriendefinition* „das Auswertungsinstrument durch theoretische Überlegungen [bestimmt] ...[,] [a]us Voruntersuchungen, aus dem bisherigen Forschungsstand, aus neu entwickelten Theorien oder Theorienkonzepten“ (ebd.). Begründet werden kann dieses Vorgehen dadurch, dass die interessierenden Kompetenzfacetten der (angehenden) Lehrpersonen bei der Entwicklung des zugrunde liegenden Kompetenzstrukturmodells – auf Basis vorliegender wissenschaftlicher Theorien und Erkenntnisse sowie durch weiterführende theoretische Überlegungen – in Anlehnung an das Vorgehen bei COACTIV „systematisch analysiert ... [und] in Bezug gesetzt [wurden] zur Gestaltung und Qualität des Unterrichts und zu den Leistungen der Schüler/-innen“ (Weschenfelder, 2014a, S. 66). Die Kriterien qualitativ hochwertigen Unterrichts

tens können also, da sie Annahmen *vor bzw. während* der Modellierung waren, nicht ex post definiert werden. Darüber hinaus wird in der vorliegenden Arbeit zwar davon ausgegangen, dass das prozedurale fachdidaktische Wissen durch zielgerichtete Übungen aufgebaut werden kann, um sich im Unterricht in Form ‚flüssigen Handelns‘ zu zeigen. Da jedoch in dieser Studie zunächst ausschließlich Lehramtsstudierende befragt werden, die im Gegensatz zu erfahrenen Lehrkräften über vergleichsweise wenig ausgeprägtes Erfahrungswissen verfügen dürften, kann das von ihnen explizierte Wissen nicht als ein verallgemeinerbares Kriterium qualitativ hochwertiger Rückmeldungen angesehen werden. Zwar wurden die vorliegenden Reaktionen der Testpersonen für die Weiterentwicklung und Verfeinerung der Ratingkategorien ergänzend gesichtet, jedoch wären für ein induktives Vorgehen auch Lehrpersonen mit langjähriger Berufserfahrung zu befragen, um dann aus dem zur Verfügung stehenden Antwortmaterial verallgemeinerbare Kategorien ableiten zu können.

Die Analyse der Merkmale hoher Qualität des Unterrichtens unter Berücksichtigung der in Kapitel 2.3.3.3 eingenommenen domänenspezifischen Perspektive lässt die Bildung folgender Kategorien zur Beurteilung der Qualität von mündlichen Rückmeldungen durch die Probanden zu: *fachliche Korrektheit*, *kognitive Aktivierung*, *Orientierung am (Fehl-)Verständnis der Lernenden*³⁸, *Klarheit und Strukturiertheit* sowie *demokratische Unterrichtskommunikation*. Um den Ratern die Beurteilung der Qualität der Rückmeldungen zu ermöglichen, wird ein Ratingmanual verwendet, das nach Mayring (2010) für jede der fünf Ratingkategorien zunächst (1) eine genaue Definition enthält, auf deren Basis (2) Ankerbeispiele gegeben werden, die die Zuordnung zur jeweiligen Kategorie erleichtern sollen, und das schließlich (3) spezifische Kodierregeln für die beiden Aufgabentypen erläutert, sofern diese für die jeweilige Kategorie relevant sind (D. Neumann, 2015, S. 32ff.).

In Expertenworkshops wurden die Kategorien erprobt und weiter ausdifferenziert. Die qualitative Beurteilung der Rückmeldungen erfolgt entlang eines von Lindmeier (2011) vorgeschlagenen Entscheidungsbaumes, der die drei Urteilkategorien „*Adequate Response*“ (S. 146) (*adäquate Rückmeldung*), „*Partially Adequate Response*“ (ebd.) (*teilweise adäquate Rückmeldung*) sowie „*Not Adequate Response*“ (ebd.) (*inadäquate Rückmeldung*) beinhaltet, wobei Lindmeier (ebd., S. 145) diese drei Hauptkategorien nochmals unterteilt und dadurch

³⁸ Hiermit angesprochen ist das Qualitätsmerkmal *Schülerorientierung*, das jedoch, um Verwechslung mit dem *didaktischen Prinzip* der Schülerorientierung (vgl. Kapitel 2.3.3.3) zu vermeiden, umbenannt wird.

neun Unterkategorien erhält. Eine solche detaillierte Unterteilung hat sich in den Expertenworkshops für die vorliegende Arbeit als nicht praktikabel erwiesen, weshalb hier eine Beurteilung der Qualität der Rückmeldungen in dem entsprechend ‚gröberen‘ Urteilsschema erfolgt (D. Neumann, 2015, S. 53).

Wegen der unterschiedlichen Charakteristika der beiden Antworttypen *Erklärung* und *unterstützender Impuls* sind jeweils spezifische Kodierregeln anzuwenden (ebd., S. 148ff.). So kann beim Antworttyp *Erklärung* Folgendes vorausgesetzt werden: „Expert teachers are expected to be better in identifying the relevant aspects of instructional situations, to act according to a certain goal, and should thus be able to provide clear feedback“ (ebd., S. 149). Es werden hierbei die Kategorien der fachlichen Korrektheit sowie der Klarheit und Strukturiertheit hervorgehoben (ebd.). Die nachfolgende Rückmeldung, die im Ratingmanual als Beispiel für eine durchgängig adäquate Rückmeldung gegeben wird, lässt die qualitativen Merkmale einer hochwertigen Erklärung als Reaktion auf einen Schülerfehler erkennen (der Schüler hat zuvor in einer Karikatur den Parteinaamen *NPD* ausformuliert als *Nationalsozialistische Partei Deutschlands*):

„Der Parteiname *NPD* steht nicht, wie du vermutet hast, für *Nationalsozialistische Partei Deutschlands*, sondern für *Nationaldemokratische Partei Deutschlands*. Wie wir schon besprochen haben, verfolgt die NPD eine sogenannte Legalitätstaktik. Das heißt, sie will die Demokratie mit ihren eigenen Mitteln schlagen. Der Begriff *nationalsozialistisch* würde eine zu große Ähnlichkeit mit der verbotenen ehemaligen Partei Adolf Hitlers, der *Nationalsozialistischen Deutschen Arbeiterpartei* aufweisen. Mit dem Wort *demokratisch* in ihrem Namen gibt die NPD vor, sich an demokratische Verfassungsprinzipien zu halten. Ob sie das tatsächlich tut, wird aber politisch und öffentlich diskutiert.“ (D. Neumann, 2015, S. 46f.)

Diese Rückmeldung ist fachlich korrekt und zielt auf die Induktion konzeptuellen Verständnisses bei dem Schüler ab. Sie ist darüber hinaus klar strukturiert und lokalisiert den Schülerfehler genau (Lindmeier, 2011, S. 148); die eigene Urteilsbildung des Schülers wird nicht behindert. Der durchgängige Bezug auf den/die Schüler/-in im Sinne der Kategorie *Orientierung am (Fehl-)Verständnis der Lernenden* lässt den ko-konstruktiven Charakter einer Erklärung erkennen (‚explaining with him‘ statt ‚explaining to him‘ nach van de Sande und Greeno, 2010, S. 69).

Die Kategorie der Klarheit und Strukturiertheit wird demgegenüber in den Urteilskategorien für die Qualität eines *unterstützenden Impulses* weniger deutlich betont (Lindmeier, 2011, S. 152; D. Neumann, 2015, S. 56), da davon ausgegangen wird, dass es durch die Vorgabe, die

Lösung bei diesem Aufgabentyp nicht zu verraten, schwieriger ist, mit einer klaren und strukturierten Formulierung zu reagieren. Für beide Antworttypen gilt, dass Unterschiede in den Kategorien *fachliche Korrektheit* und *konzeptuelles Verständnis* die wesentlichen Kriterien der Unterscheidung sind, um zwischen einer (teilweise) adäquaten und inadäquaten Antwort zu unterscheiden. Da sich die Schüler/-innen teilweise unpräzise äußern, sind die Spezifika in der Kategorie *Orientierung am (Fehl-)Verständnis der Lernenden* in den Unterschieden zwischen den Niveaustufen weniger deutlich zu erkennen (D. Neumann, 2015, S. 57).³⁹

6.3.2.2 Ratingverfahren

Das Ratingverfahren in der vorliegenden Arbeit kann als eine Mischform aus *analytischem Rating* und *holistischem Rating* (Klein et al., 1998) beschrieben werden. Bei einem holistischen Rating wird ein Globalurteil zu gegebenen Antworten gebildet (Bos, Klieme & Köller, 2010), während bei einem analytischen Rating alle Bewertungskriterien einzeln angewendet werden, insofern die „Unterskalen ... getrennt bewertet und beurteilt werden“ (Tschirner, 2001, S. 19). Zwar werden die Antworten in der vorliegenden Arbeit anhand der Ratingkategorien nach unterschiedlichen Kriterien bewertet, jedoch kann – was Voraussetzung für ein analytisches Rating wäre – nicht davon ausgegangen werden, dass die zu beurteilenden Aspekte in den Antworten der Probanden voneinander unabhängig sind. So wird die Bewertung der fachlichen Korrektheit einer Antwort analytisch nur schwer trennbar sein von der Frage, ob die Antwort auf die Induktion von Wissen auf konzeptuellem Niveau bei den Lernenden abzielt (Lindmeier, 2011, S. 141). Die Mischform des Ratingverfahrens zeichnet sich in Anlehnung an das Vorgehen bei Lindmeier daher durch Folgendes aus: „The ratings relied on a holistic judgement of trained raters, but different analytic criteria were implemented in the rating guidelines“ (ebd.).

Mit dem Begriff der *trained raters* ist angesprochen, dass die Qualität der Rückmeldungen durch zwei geschulte Personen in Form eines Expertenratings (*expert scoring*) bewertet wird. In diesem Verfahren kann zwar dem Problem des Rater-Bias nicht abschließend be-

³⁹ Hiermit wird auch dem Umstand Rechnung getragen, dass in der vorliegenden Studie mit Videovignetten aus authentischen Unterrichtssituationen gearbeitet wird, die in ihrer Darstellung der Verständnisprobleme der Schüler/-innen weniger eindeutig sind als vergleichbare nachgespielte Unterrichtsausschnitte.

gegnet werden (Praetorius, 2014, S. 49f.), jedoch wird die Objektivität der Ratings durch voneinander unabhängige Personen im Gegensatz zu einem Rating durch eine Einzelperson deutlich erhöht (Bortz & Döring, 2006, S. 185). Die Rater wurden auf Basis des vorliegenden Manuals in Expertensitzungen zur Bewertung der offenen Antworten in den hier interessierenden Aspekten geschult.⁴⁰ Im ersten Anlauf des Auswertungsverfahrens zeigten sich zunächst Schwierigkeiten, da die Kategorie *kognitive Aktivierung* von den Ratern offenbar unterschiedlich ausgelegt wurde; eine ähnliche Problematik in einer inhaltsähnlichen Ratingkategorie berichten Rehm und Bölsterli (2014, S. 221). Das Ratingmanual wurde daher in dieser Kategorie nochmals theoretisch spezifiziert und es wurden weiteren Ankerbeispiele angegeben, die den Ratern die Beurteilung erleichtern sollten. Auf Grundlage des überarbeiteten Ratingmanuals fand dann eine Nachschulung statt (vgl. auch ebd.).

Die Qualität der Antworten wird anhand der insgesamt 500 Transkriptionen (D. Neumann, 2015, S. 84ff.) der mündlichen Rückmeldungen der Testpersonen beurteilt, wobei die Audioaufzeichnungen z. B. bei mehrdeutigen Antworten hinzugezogen werden können. Auch die Äußerungen der Schüler/-innen liegen den Ratern bei Bedarf zum besseren Verständnis des Antwortverhaltens vor (ebd., S. 70ff.). Sie sind aufgefordert, die Videovignette vor der Bewertung stets mindestens einmalig anzuschauen. Die Transkription erfolgte anhand der Audiospur der vorliegenden Videodateien mit den Antworten der Probanden. Bei der Verschriftlichung von Audiomaterial sind vorab Entscheidungen hinsichtlich der Transkriptionsdarstellung zu treffen, wobei Entscheidungsaspekte gegeneinander abgewogen werden müssen. So werden Transkripte eingesetzt, um „Ergebnisse anhand einer naturalistischen, passiv registrierenden Datengrundlage gewinnen“ (Deppermann, 2008, S. 46) zu können, was einerseits eine möglichst genaue Darstellung der sprachlichen Äußerungen erfordert, andererseits ist aber auch der Verwendungszweck der Transkripte zu berücksichtigen (ebd.). Im Hinblick auf den für diese Arbeit relevanten Zweck der Transkripte wurden aus Gründen der Praktikabilität und Lesbarkeit beispielsweise prosodische Merkmale wie die Lautstärke und nonvokale Phänomene (ebd., S. 43ff.) nicht transkribiert. Dialektbedingte sowie allgemeine sprachliche Besonderheiten und Fehler wurden in die korrekte schriftsprachliche Form überführt (z. B. Wortzusammensetzungen wie ‚kannste‘ in ‚kannst du‘); die Interpunktion in den Transkripten erfolgte sinngemäß. Grundlegende Grammatik- oder Tempusfehler sowie unkorrekte (Fach-)Begriffe wurden nicht korrigiert. Akzentuierungen von Wörtern oder Wortteilen wur-

⁴⁰ Für eine Problematisierung ggf. fehlender Effektivität von Rater-Trainings vgl. Praetorius (2014, S. 76).

den kursiv gesetzt (,Das ist zwar grundsätzlich richtig, *aber* du musst auch bedenken ...‘). Füllwörter (,äh‘ oder ,ähm‘) wurden transkribiert, ebenso Wort- bzw. Satzabbrüche. Letztere sind mit einem Bindestrich (z. B. ,ausges-ausgefallen‘) gekennzeichnet. Sprechpausen wurden transkribiert, sofern sie mindestens drei Sekunden andauerten (im Transkript z. B. [3 Sek. Pause]). Unverständliche Wörter wurden mit der Bemerkung ,[unverständlich]‘ gekennzeichnet. Eigene Anmerkungen des Transkribierenden wurden in eckige Klammern gesetzt, bspw. ,[Anfang Metakommunikation des Probanden]‘ und ,[Ende Metakommunikation des Probanden]‘ oder ,[Abbruch]‘. Abbildung 23 zeigt ein entsprechendes Beispieltranskript.

#00:00:44-9#

Ja, genau. Du-Du stellst auf jeden Fall was Wichtiges heraus. Es ist auf jeden Fall in einer Demokratie wichtig, dass alle-dass möglichst viele verschiedene Meinungen gehört werden, dass ähm dass sie ernst genommen werden und dass wirklich jeder das Gefühl hat, er kann sagen, was er möchte. Recht auf freie Meinungsäußerung. Es ist ja auch nicht so, dass *verfassungsfeindliche* Äußerungen per se verboten sind. Der Unterschied besteht zwischen *verfassungsfeindlich* und *verfassungswidrig*. Insofern eine Partei verfassungsfeindlich ist, hat das Bundesverfassungsgericht noch nicht die Möglichkeit, sie direkt zu verbieten. Vielmehr geht es darum, ob die Aktionen, die wirklich von der Partei durchgeführt werden, verfassungswidrig sind. Das bedeutet, wenn sie eben auf die Zerstörung unserer Grundordnung eingehen und auf die gewaltsame n-Durchsetzung von Zielen, die mit unserer äh demokratisch freiheitlich demokratischen Grundordnung nicht vereinbar sind.

#00:01:36-1#

Abbildung 23: Beispieltranskript einer Rückmeldung eines Probanden (Prob_ID 3) in der Skala zum prozeduralen politikdidaktischen Wissen (Antworttyp *Erklärung*)

7 Ergebnisse der Studie

Dieses Kapitel gibt einen Überblick über die empirischen Befunde der Studie. Zunächst wird die Stichprobe beschrieben (Kapitel 7.1) und mögliche Unterschiede zwischen Gruppen werden mithilfe explorativer Signifikanztests untersucht. Anschließend wird entlang der in Kapitel 5 formulierten Forschungsfragen die Item- und Skalenkonstruktion analysiert und es werden die Testleistungen der Probanden untersucht. Schließlich werden die strukturellen und differenziellen Annahmen zur Modellierung auf Grundlage explorativer quantitativer Datenanalysen erkundet (Kapitel 7.2) und kritisch reflektiert (Kapitel 7.3).

7.1 Stichprobe

An der Studie nahmen angehende Lehrkräfte für politisch-sozialwissenschaftliche Unterrichtsfächer teil (50 Studierende des Lehramts Sozialwissenschaften der Universität Duisburg-Essen). Eine Probandin hatte das Studium zum Zeitpunkt der Befragung bereits abgeschlossen und befand sich im Vorbereitungsdienst. Ihre Daten wurden in die Auswertung mit einbezogen, da zum einen die Studie nicht der Generierung verallgemeinerbarer Aussagen (Generalisierbarkeit) dienen soll und sie somit auch nicht auf eine repräsentative Stichprobe angewiesen ist. Zum anderen befand sich die Lehramtsanwärterin – wie ein nach der Befragung durchgeführtes Interview ergab – zum Zeitpunkt der Befragung erst seit wenigen Wochen im Vorbereitungsdienst, sodass davon ausgegangen werden kann, dass sie sich in den hier interessierenden Merkmalen (insbesondere in dem für das prozedurale fachdidaktische Wissen bedeutsamen unterrichtlichen Erfahrungswissen) nicht wesentlich von den übrigen Versuchspersonen unterscheidet.

Die Teilnahme an der Studie war für die Studierenden freiwillig, insofern handelt es sich bei der Stichprobe um eine Gelegenheitsstichprobe. Die freiwilligen Testpersonen einer solchen nichtprobabilistischen Stichprobe entsprechen nicht dem Durchschnitt der Gesamtpopulation und können sich von den Personen, die nicht teilnahmen, in Eigenschaften wie einer besseren schulischen Ausbildung, höherer Intelligenz oder einem höheren Bedürfnis nach sozialer Anerkennung unterscheiden (Bortz & Döring, 2006, S. 704). Dies ist bei der Diskussion der Befunde der Studie zu berücksichtigen.

Zwar studierten alle Teilnehmenden das gleiche Studienfach *Sozialwissenschaften* als Erst- oder Zweitfach mit einem weiteren Unterrichtsfach und bildungs- und erziehungswissenschaftlichen Studienanteilen, jedoch unterscheidet sich die Studienstruktur in den an der Universität Duisburg-Essen zum Befragungszeitpunkt bereits auslaufenden Lehramtsstudiengängen gem. Lehrerausbildungsgesetz (LABG) 2002 (24 Probanden) teilweise von den neuen Studiengängen gem. LABG 2009 (26 Probanden), unter anderem hinsichtlich des stärker modularisierten Studienverlaufs sowie in Aufbau und Struktur der schul- bzw. berufspraktischen Anteile in Form von Praktika in der konsekutiven Bachelor-/Master-Studienstruktur.⁴¹ Es sollen daher – wo für die anschließende Interpretation der vorliegenden Daten relevant – in den folgenden Beschreibungen der Stichprobe auch Unterschiede zwischen den beiden Studierendengruppen nach Studienabschluss fokussiert werden, um schlussfolgern zu können, ob die anschließenden statistischen Verfahren getrennt nach Subgruppen erfolgen müssen oder ob die beiden Studierendengruppen zu einer Gesamtgruppe klassiert werden können.

Von den insgesamt teilnehmenden 50 Probanden sind 29 weiblich und 21 männlich. Die Probanden waren zum Zeitpunkt der Erhebung zwischen 21 und 30 Jahren alt; das Durchschnittsalter betrug aufgerundet 24 Jahre ($M = 23.92$, $SD = 2.25$). Die relativ niedrige Streuung in den Angaben zum Alter (30 % der Befragten waren 22 Jahre alt; das Alter von 74 % der Befragten lag in einer Spannweite zwischen 22 und 25 Jahren) lässt sich vor allem durch einen Kohorteneffekt im Bachelor-Studiensystem (52 % aller Befragten) bezogen auf die Einführung der konsekutiven Bachelor-Lehramtsstudiengänge an der Universität Duisburg-Essen zum Wintersemester 2011/12 in Verbindung mit dem Umstand erklären, dass zum Befragungszeitpunkt eben diese Kohorte im Studienverlauf erstmalig politikdidaktische Inhalte studiert hatte oder studierte und somit für eine Teilnahme an der vorliegenden Studie zum fachdidaktischen Wissen und Können infrage kam; vgl. Tabelle A. 2 im Anhang für weitere Details.

⁴¹ Ein wesentlicher konzeptioneller Unterschied zwischen den Studiengängen besteht darin, dass der grundlegende Bachelor-Studiengang den Absolventinnen und Absolventen einen ersten berufsqualifizierenden Abschluss ausweist, der zwar noch nicht für den Zugang zum Vorbereitungsdienst befähigt, jedoch unter anderem – zumindest die Anzahl betreffend – umfangreichere Praxisphasen als die Studiengänge gem. LABG 2002 beinhaltet; für eine kritische Auseinandersetzung mit den Möglichkeiten und Grenzen der reformierten Praxisphasen in der ersten Phase der Lehrerbildung vgl. Offenberg & Walke (2013).

In Annahme der Nullhypothese H_0 , dass zwischen den Probanden mit dem Abschluss Bachelor of Arts und den Probanden mit dem Abschluss Staatsexamen kein signifikanter Unterschied in der Merkmalsausprägung *Geschlecht* besteht, weist der Chi-Quadrat-Test für die asymptotische Signifikanz (zweiseitig; Kontinuitätskorrektur) einen Wert von $\chi^2 = .81$ aus; dieser Wert liegt über dem Signifikanzniveau von $\alpha = .05$. Es zeigt sich daher kein signifikanter statistischer Zusammenhang zwischen den Merkmalsausprägungen. Die Nullhypothese wird behalten; es kann davon ausgegangen werden, dass zwischen den beiden Gruppen kein signifikanter Unterschied in der Merkmalsausprägung *Geschlecht* besteht. Die Annahme der Nullhypothese H_0 , dass zwischen den beiden Gruppen kein statistisch signifikanter Unterschied in der Merkmalsausprägung *Alter* besteht, wird mithilfe des U-Tests von Mann-Whitney überprüft.⁴² Der Signifikanzwert (asymptotische Signifikanz aufgrund der Größen der Teile der Gesamtstichprobe von $n_1 > 10$ und $n_2 > 10$) liegt mit $p < .001$ deutlich unter dem Signifikanzniveau von $\alpha = .05$, weshalb die Nullhypothese verworfen wird. Es kann davon ausgegangen werden, dass ein statistisch signifikanter Unterschied in der Variable *Alter* zwischen den Probanden mit den Abschlüssen Bachelor of Arts und Staatsexamen besteht. Der Ausprägungsgrad unterrichtspraktischen Erfahrungswissens (Praktika, [neben-]berufliche oder sonstige außeruniversitäre Tätigkeiten) bei den Probanden wird im späteren Verlauf dieses Kapitels noch dargestellt. Jedoch kann zur Beantwortung der Frage, ob ein Zusammenhang zwischen dem Alter und dem unterrichtspraktischen Erfahrungswissen der Probanden feststellbar ist, der wiederum hinsichtlich des dargestellten signifikanten Unterschieds im durchschnittlichen Alter zwischen den Gruppen zu diskutieren wäre, für die vorliegenden Daten festgehalten werden, dass diese auf keinen signifikanten Zusammenhang zwischen dem Alter und dem unterrichtspraktischen Erfahrungswissen hindeuten: Die getrennte Berechnung des Rangkorrelationskoeffizienten von Spearman (ρ) für beide Gruppen für die Variablen *Alter* und *lehr-/unterrichtspraktische Erfahrungen in Zeitstunden* zeigt für die Bachelor-Studierenden einen Korrelationskoeffizienten von $\rho = .02$, der mit $p = .91$ deutlich über dem Signifikanzniveau von $\alpha = .05$ liegt, und für die Studierenden mit dem Ab-

⁴² Parametrische Verfahren, wie zum Beispiel der t-Test, können bei dieser Variable nicht zur Anwendung kommen, da die Daten zur Variable *Alter* keiner Normalverteilung folgen. Der Shapiro-Wilk-Test auf Normalverteilung zeigt für diese Variable für die Studierendengruppe mit dem Abschluss Bachelor of Arts einen Signifikanzwert von $p < .001$; da der p -Wert also deutlich kleiner als das Signifikanzniveau von $\alpha = .05$ ist, wird die Nullhypothese, dass die Daten normalverteilt sind, verworfen. Der Signifikanzwert der Studierendengruppe mit dem Abschluss Staatsexamen liegt bei $p = .03$; auch hier liegt der p -Wert deutlich unter dem Signifikanzniveau von $\alpha = .05$, weshalb die Nullhypothese einer Normalverteilung auch hier verworfen wird.

schluss Staatsexamen einen Korrelationskoeffizienten von $\rho = .23$ ($p = .29$), der auf einen lediglich schwachen linearen Zusammenhang zwischen den beiden Variablen hindeutet.

Von den Befragten mit dem anvisierten Studienabschluss Erste Staatsprüfung belegten 5 den Studiengang für die Schulformen Haupt-, Real- und Gesamtschule (HRGe; Sekundarstufe I) und 19 den Studiengang für die Schulform Gymnasium sowie die entsprechenden Jahrgangsstufen der Gesamtschule (Gy/Ge; Sekundarstufe II). Von den Probanden im Bachelor-Studiengang studierten 7 Befragte für die Schulformen der Sekundarstufe I und 19 Befragte für die Schulformen der Sekundarstufe II; für eine Zusammenschau der relevanten Daten der Stichprobe zu Studienabschluss, Schulform und Studienverlauf (Fachsemester) vgl. Tabelle A. 3 im Anhang. Auf mögliche Unterschiede zwischen den Probanden gruppiert nach angestrebtem Studienabschluss hinsichtlich der späteren Schulform (HRGe/GyGe) soll in diesem Zusammenhang nicht näher eingegangen werden, da sich die Studienverläufe innerhalb der Studiengänge – bis auf einen Unterschied im Workload zwischen HRGe und GyGe in den Bachelor-Studiengängen, der jedoch nur einen Leistungspunkt umfasst und im Studium in der Regel lediglich durch einen leicht angehobenen Zeitumfang der Prüfungsleistung (Referat) in der relevanten Lehrveranstaltung eingefordert wird – nicht voneinander unterscheiden (UDE, 2014a, 2014b).

Danach befragt, welche Lehrveranstaltungen mit politikdidaktischen Inhalten sie zum Befragungszeitpunkt bereits studiert hatten, gaben von den 26 Studierenden des Bachelor-Lehramtes 23 Studierende an, mindestens die Veranstaltung (Vorlesung mit abschließender Klausur) *Grundlagen der Fachdidaktik der Sozialwissenschaften* absolviert zu haben; ein Proband hatte diese Veranstaltung noch nicht belegt, studierte jedoch zum Zeitpunkt der Befragung bereits ein politikdidaktisches Vertiefungsseminar. Die beiden fehlenden Werte lassen sich in einem Fall auf ein technisches Problem zurückführen, da hier die Antwortzeit der Probandin im Computerprogramm bei null Sekunden lag und daher davon ausgegangen werden kann, dass die Frage unabsichtlich übersprungen wurde; im anderen Fall wurden keine Angaben gemacht (vgl. Tabelle A. 4 im Anhang). Von den Studierenden mit dem Studienabschluss Erste Staatsprüfung gaben ebenfalls 23 Studierende an, mindestens die Grundlagenveranstaltung im Grundstudium absolviert zu haben; 18 der insgesamt 24 Probanden hatten bereits alle vorgesehenen Lehrveranstaltungen mit politikdidaktischen Inhalten ihres Studiengangs (neben der Grundlagenvorlesung zwei weitere Vertiefungsseminare im Haupt-

studium) erfolgreich besucht (vgl. Tabelle A. 5 im Anhang). Unter Ausschluss der fehlenden Werte zeigt sich zwischen den Probanden gruppiert nach Studienabschluss eine vollständige Übereinstimmung in der Merkmalsausprägung *Grundlagenvorlesung besucht*; die Verteilung der Merkmalsausprägung ist in beiden Gruppen gleich. Zusammengefasst für beide Studiengruppen lässt sich feststellen, dass mit den aufgeführten Ausnahmen 92 % aller Befragten mindestens die einführende Grundlagenveranstaltung sowie 78 % aller Befragten darüber hinaus bereits mindestens eine vertiefende politikdidaktische Lehrveranstaltung besucht hatten (vgl. Tabelle A. 6 im Anhang) und somit über mindestens basales deklaratives Wissen zur Politikdidaktik verfügen haben dürften.⁴³

Auch in einer weiteren Variable, die auf Unterschiede in der Befragung zwischen den Probanden mit dem Abschluss Erste Staatsprüfung und den Probanden in den Bachelor-Studiengängen hindeuten könnte, den studienbegleitend zu absolvierenden Praktika, zeigen sich deutliche Ähnlichkeiten zwischen beiden Gruppen. Von den Bachelor-Studierenden hatten zum Befragungszeitpunkt alle Probanden bereits erste Praktikumserfahrungen (mindestens Eignungspraktikum) gesammelt; 20 der 26 Probanden hatten zudem alle im Bachelor-Studiengang vorgesehenen Praktika (Eignungspraktikum, Orientierungspraktikum sowie Berufsfeldpraktikum⁴⁴) absolviert (vgl. Tabelle A. 7 im Anhang). Alle Studierenden mit Abschluss Erste Staatsprüfung hatten zum Zeitpunkt der Befragung mindestens das im Grundstudium zu absolvierende Orientierungspraktikum abgeschlossen; 18 der 24 Studierenden gaben an, auch das im Hauptstudium verortete Fachpraktikum für das Lehramt Sozialwissenschaften bereits abgeleistet zu haben (vgl. Tabelle A. 9 im Anhang).

Trotz der konzeptionellen Unterschiede in den Praktika zwischen den unterschiedlichen Lehramtsstudiengängen zeigen sich hinsichtlich des Umfangs wiederum deutliche Ähnlich-

⁴³ Hierbei wird auf Basis der bildungsadministrativ vorgegebenen inhaltlichen Anforderungen an die fachdidaktischen Studienbestandteile sowie die hiermit verbundenen Modulprüfungen angenommen, dass für das erfolgreiche Bestehen (absolute untere Bestehensgrenze: Note 4,0) der standardisierten Modulabschlussprüfung (Klausur zur Vorlesung *Grundlagen der Fachdidaktik der Sozialwissenschaften*) bei den Studierenden deklaratives politikdidaktisches Wissen auf mindestens ausreichendem Niveau vorhanden sein muss.

⁴⁴ Das Berufsfeldpraktikum (BFP) kann wahlweise in einer Schule oder in einer außerschulischen bildungsorientierten Einrichtung absolviert werden. Von den 24 Bachelor-Studierenden, die ihr Berufsfeldpraktikum zum Befragungszeit bereits abgeschlossen hatten, gaben 19 Studierende an, ihr BFP an einer Schule absolviert zu haben; 4 Studierende hatten ihr BFP an einer außerschulischen bildungsorientierten Einrichtung abgeleistet (ein fehlender Wert infolge versehentlichen Überspringens der Frage im Computerprogramm [Antwortdauer null Sekunden]; vgl. auch Tabelle A. 8 im Anhang).

keiten zwischen den beiden Studierendengruppen. In den Studiengängen gem. LABG 2002 umfassen die schulischen Praxisphasen mindestens 14 Wochen bei einem Umfang von ungefähr 20 Zeitstunden pro Woche; das Orientierungspraktikum im Grundstudium umfasst dabei 120 Stunden Schulaufenthalte und die Fachpraktika im Hauptstudium 80 Stunden je Unterrichtsfach (UDE, 2008). In den Bachelor-Lehramtsstudiengängen umfasst das Eignungspraktikum 20 Tage Schulaufenthalte und das Orientierungspraktikum sowie das Berufsfeldpraktikum jeweils 80 Zeitstunden Schulaufenthalte (ZLB, 2014). Von den Befragten im Lehramt gem. LABG 2002 haben demnach 75 % der Studierenden Schulaufenthalte im Umfang von mindestens 200 Zeitstunden absolviert; von den Bachelor-Studierenden haben bei einer vergleichbaren Umrechnung der Schulaufenthalte im Orientierungspraktikum⁴⁵ 77 % der Befragten Praktikumsaufenthalte im Umfang von mindestens 240 Zeitstunden abgeleistet.⁴⁶ Somit zeigt sich hinsichtlich der durch die Studien- bzw. Prüfungsordnungen vorgeschriebenen Praxisphasen, dass zwischen den Studierenden der Lehramtsstudiengänge keine wesentlichen Unterschiede im Umfang der tatsächlich absolvierten Praktikumsaufenthalte zu erwarten sind, zumindest was die untere Stundengrenze betrifft.

Die in den Praktika gewonnenen Erfahrungen sind hierbei nicht vollumfänglich mit unterrichtspraktischen Tätigkeiten gleichzusetzen. Ziele der Praktika sind neben der Erprobung eigenen – stets durch Lehrkräfte begleiteten – Unterrichts das theoriegeleitete Analysieren, Beobachten und Reflektieren schulischer sowie (fach-)unterrichtlicher Prozesse und Strukturen (MSW, 2014b, S. 1). In den Bachelor-Studiengängen ist die Umsetzung späterer berufspraktischer Tätigkeiten wie das eigenständige – jedoch weiterhin durch erfahrene Lehrkräfte unterstützte – Unterrichten in der Schule schwerpunktmäßig in die Master-Studiengänge verlagert, hier in das Praxissemester.⁴⁷ Da also von nur begrenztem unterrichtspraktischem

⁴⁵ Dabei entsprechen 20 Tage etwa vier Arbeitswochen, was einen Gesamtumfang von etwa 80 Stunden bei 20 Zeitstunden in der Woche ergibt.

⁴⁶ Es ist zu berücksichtigen, dass wegen der Wahlmöglichkeiten nicht alle Studierenden ihr Berufsfeldpraktikum an einer Schule geleistet haben. Die Mehrzahl – insgesamt 19 der 26 Studierenden – gaben jedoch an, ihr BFP an einer Schule absolviert zu haben.

Zudem enthalten die meisten fachspezifischen Vorgaben für die Berufsfeldpraktika an außerschulischen bildungsorientierten Einrichtungen Regelungen hinsichtlich der möglichen Praktikumsorte (beispielsweise zu einer ausreichenden Fachspezifität der Praktikumsaktivitäten und dazu, dass die Praktikanten lehrende Tätigkeiten ausüben können müssen), die vergleichbare Rahmenbedingungen schulischer sowie außerschulischer Praktikumsorte schaffen.

⁴⁷ Das Praxissemester umfasst Schulaufenthalte von insgesamt fünf Monaten.

Erfahrungswissen der Studierenden ausgegangen werden kann, das sie in den studienbegleitenden Praktika erworben haben, wurden die Probanden auch nach ihren außeruniversitären beruflichen Erfahrungen befragt, in denen sie lehrende Tätigkeiten ausübten. Von den 24 Befragten im Lehramt mit dem Abschluss Staatsexamen gaben 19 Probanden an, zum Zeitpunkt der Studie neben den Praktikumstätigkeiten auch außeruniversitäre praktische Lehr- bzw. Unterrichtstätigkeiten – beispielsweise in Form von Nachhilfeunterricht – ausgeübt zu haben. Eine Teilnehmerin gab an, unterrichtliche Erfahrungen ausschließlich in Form außeruniversitärer (neben-)beruflicher Tätigkeiten gesammelt zu haben. Drei Probanden mit dem angestrebten Abschluss Erste Staatsprüfung hatten zum Befragungszeitpunkt nur Praktika absolviert und keine außeruniversitären Lehr-/Unterrichtserfahrungen gesammelt; ein Proband machte hierzu keine Angaben. Von den 26 Bachelor-Studierenden hatten 19 Studierende neben den Praktikumstätigkeiten auch außeruniversitäre Lehr-/Unterrichtstätigkeiten ausgeübt; die übrigen Studierenden hatten zum Zeitpunkt der Befragung lediglich universitäre Praktikumserfahrungen gesammelt.

In Annahme der Nullhypothese H_0 , dass zwischen den Studierenden mit dem Abschluss Bachelor of Arts und den Studierenden mit dem Abschluss Staatsexamen kein signifikanter Unterschied in der Merkmalsausprägung *außeruniversitäre Lehr-/Unterrichtserfahrung* (nur *außeruniversitäre Erfahrungen* und *außeruniversitäre sowie Praktikumserfahrungen*) besteht, weist der Chi-Quadrat-Test für die asymptotische Signifikanz (2-seitig; Kontinuitätskorrektur) einen Wert von $p = .40$ aus. Dieser Wert liegt über dem Signifikanzniveau von $\alpha = .05$, es zeigt sich somit kein statistisch signifikanter Zusammenhang zwischen den Merkmalsausprägungen; die Nullhypothese wird behalten, weshalb auch in Bezug auf die außeruniversitären Lehr-/Unterrichtserfahrungen davon ausgegangen werden kann, dass keine signifikanten Unterschiede zwischen den Gruppen vorliegen.

Schließlich wurden die Probanden im Zusammenhang mit den bisherigen unterrichtspraktischen Tätigkeiten danach befragt, in welchem Umfang (Zeitstunden) sie zum Zeitpunkt der Befragung Lehr-/Unterrichtserfahrungen gesammelt hatten (offenes Antwortformat). Die Spannweite der Antworten lag zwischen null Zeitstunden (keine unterrichtspraktischen Erfahrungen) und 750 Zeitstunden, wobei die Streuung der Angaben groß ist, was sich dadurch erklären lässt, dass sich die Frage sowohl auf die Praktika bezog als auch auf die eigenständigen außeruniversitären Tätigkeiten, die zum Zeitpunkt der Befragung bereits ausgeübt

wurden. Aufgrund der Sensibilität für Extremwerte und der in Abbildung A. 1 im Anhang bereits erkennbaren Ausreißer eignet sich die Bestimmung des arithmetischen Mittels zur Angabe der durchschnittlichen Lehr-/Unterrichtserfahrungen nicht. Demgegenüber kann der Median (\tilde{x}) wegen seiner Robustheit gegenüber Ausreißern für die Bestimmung des Mittelwerts der Verteilung als aussagekräftiger angesehen werden. Nach Studienabschluss getrennt berechnet liegt der Median der angegebenen Lehr-/Unterrichtserfahrung für die Studierenden mit dem Abschluss Bachelor of Arts bei $\tilde{x} = 9$ Zeitstunden und für die Studierenden mit dem Abschluss Erste Staatsprüfung (Staatsexamen) bei $\tilde{x} = 14.5$ Zeitstunden.

Eine andere Möglichkeit zur Bestimmung des Mittelwerts bei Vorliegen von Extremwerten besteht in der Berechnung des Alpha-getrimmten arithmetischen Mittelwerts (\bar{x}_α), dem „aus den sortierten Werten vom unteren und oberen Ende Werte entfernt [werden]“ (Kohn & Öztürk, 2013, S. 37), üblicherweise 5 % der Randwerte (ebd.). Der um 5 % getrimmte arithmetische Mittelwert für die angegebenen Lehr-/Unterrichtserfahrungen in Zeitstunden der Studierenden mit dem Abschluss Bachelor of Arts liegt bei $\bar{x}_\alpha = 15.91$ und für die Studierenden mit dem Abschluss Erste Staatsprüfung bei $\bar{x}_\alpha = 29.81$.

Schließlich kann der Alpha-winsorisierte Mittelwert ($\bar{x}_{\omega,\alpha}$) berechnet werden. Hierbei werden im Gegensatz zum getrimmten Mittelwert aus den sortierten Werten von den Enden die Werte nicht gestutzt, sondern die Extremwerte werden am oberen und unteren Ende parallel um den nächstgrößeren bzw. -kleineren Wert ersetzt, wodurch der Stichprobenumfang erhalten bleibt. Der Alpha-Wert bezeichnet hierbei den Wert des *cutoffs* in Prozent. Das Streudiagramm (Abbildung A. 1 im Anhang) lässt erkennen, dass eine Häufung der Angaben in der Spannweite zwischen null und 50 Zeitstunden mit wenigen Ausreißern nach oben vorliegt. 50 Zeitstunden können daher als sinnvoller oberer Wert für die Stutzung angesehen werden. Die explorative Datenanalyse aus Tabelle A. 7 im Anhang zeigt für die Gruppe der Studierenden mit dem Abschluss Bachelor of Arts als Extrema zwei Werte (80 und 240 Zeitstunden), die größer als der anvisierte Wert der Stutzung am oberen Ende sind, und für die Studierenden mit dem Abschluss Erste Staatsprüfung insgesamt drei Extremwerte, die größer als 50 Zeitstunden sind (750, 200 und 120 Stunden; vgl. Tabelle A. 9 im Anhang). Eine Stutzung der Werte am oberen und unteren Ende um 15 % ergibt den gewünschten Cutoff-Wert.⁴⁸ Der

⁴⁸ Bachelor of Arts: $\frac{26}{100} \cdot 15 = 3.9$, was aufgerundet vier Werten entspricht; Staatsexamen: $\frac{24}{100} \cdot 15 = 3.6$, was aufgerundet ebenfalls vier Werten entspricht.

15 %-winsorisierte Mittelwert für die Lehr-/Unterrichtserfahrungen in Zeitstunden der Studierenden mit dem Abschluss Bachelor of Arts beträgt $\bar{x}_{\omega,\alpha} = 14.04$ und für die Studierenden mit dem Abschluss Erste Staatsprüfung $\bar{x}_{\omega,\alpha} = 21.67$.

Die Nullhypothese H_0 , dass zwischen den beiden Studierendengruppen (LABG 2002 und LABG 2009) kein statistisch signifikanter Unterschied in der Merkmalsausprägung *Lehr-/Unterrichtserfahrung in Zeitstunden* vorliegt, wird mithilfe des U-Tests überprüft.⁴⁹ Der Signifikanzwert (asymptotische Signifikanz aufgrund der Größen der Teile der Gesamtstichprobe $n_1 > 10$ und $n_2 > 10$) liegt mit $p = .48$ über dem Signifikanzniveau von $\alpha = .05$, weshalb die Nullhypothese behalten wird. Es kann davon ausgegangen werden, dass kein statistisch signifikanter Unterschied in der Variable *Lehr-/Unterrichtserfahrung in Zeitstunden* zwischen den Studierendengruppen mit den Abschlüssen Bachelor of Arts und Staatsexamen besteht.

Zusammenfassend für die Beschreibung der Stichprobe im Hinblick auf mögliche Unterschiede zwischen den beiden Studierendengruppen gem. LABG 2002 (Staatsexamen) und gem. LABG 2009 (Bachelor of Arts) lässt sich festhalten, dass sich einzig das durchschnittliche Alter der beiden Gruppen statistisch signifikant voneinander unterscheidet. Da das Alter jedoch auf statistisch nicht signifikantem Niveau mit der Variable *Lehr-/Unterrichtserfahrung in Zeitstunden* zusammenhängt, scheint dieser Unterschied zwischen den Gruppen für die nachfolgenden Überlegungen vernachlässigbar. In den weiteren erhobenen Merkmalsausprägungen wie *Geschlecht*, *Umfang politikdidaktischer Studienanteile* sowie *Rahmen bisheriger Lehr-/Unterrichtserfahrungen* und *Lehr-/Unterrichtserfahrung in Zeitstunden* lassen sich zwischen den Gruppen keine statistisch signifikanten Unterschiede feststellen. Mögliche Unterschiede aufgrund bildungsadministrativer Vorgaben zu Studienstruktur und -organisation in Variablen wie *Inhalte politikdidaktischer Studienanteile* und *Umfang universitärer Praktikumserfahrungen* können aufgrund der vergleichbaren Studien- und Prüfungsordnung der Studiengänge sowie der polyvalenten Studienstruktur des Lehramtes Sozialwissenschaften an der Universität Duisburg-Essen für diese Untersuchung ausgeschlossen werden. Für die nachfolgenden statistischen

⁴⁹ Der Shapiro-Wilks-Test auf Normalverteilung zeigt für diese Variable für beide Studierendengruppen einen Signifikanzwert von $p < .001$; da dieser Wert deutlich kleiner als das Signifikanzniveau von $\alpha = .05$ ist, wird die Nullhypothese, dass die Daten normalverteilt sind, verworfen. Insofern können parametrische Verfahren nicht zur Anwendung kommen.

Verfahren werden die Probanden beider Subgruppen daher zur Gesamtgruppe *Studierende Lehramt Sozialwissenschaften* klassiert.

7.2 Empirische Befunde der Studie

In diesem Kapitel werden zunächst Details zur Testdauer und zum Umgang mit fehlenden Werten dargestellt (Kapitel 7.2.1). Anschließend werden die Items hinsichtlich ihrer Schwierigkeit und Trennschärfe analysiert (Kapitel 7.2.3), bevor die interne Konsistenz der Skalen thematisiert wird (Kapitel 7.2.4). Nachfolgend wird die Testschwierigkeit betrachtet (Kapitel 7.2.5) und es werden Zusammenhangsmaße der Skalen erläutert (Kapitel 7.2.6), bevor die empirischen Befunde abschließend in Kapitel 7.3 zusammengefasst und kritisch vor dem Hintergrund der Testgütekriterien reflektiert werden.

7.2.1 Testdauer und fehlende Werte

Für den Durchlauf im schriftlichen Bestandteil des Tests wurde keine obere Zeitgrenze festgesetzt; die Testdauer der Probanden variierte zwischen 18 und 27 Minuten. Demgegenüber ist die maximale Bearbeitungszeit für alle Items im Computerprogramm vKid festgelegt. Neben den in Kapitel 6.3.1.2 bereits erläuterten Vorgaben für die Items zum prozeduralen politikdidaktischen Wissen wurden die Bearbeitungszeiten in pilotierenden Maßnahmen vorab erprobt. Die Bearbeitungszeiten in der Studie decken für die Fragen zu den soziodemografischen Merkmalen der Testpersonen sowie zu den Items zum deklarativen politikdidaktischen Wissen jeweils die maximalen Bearbeitungszeiten in den Probedurchläufen ab. Die maximale Gesamtbearbeitungszeit (alle Testbestandteile) lag bei allen Teilnehmenden zwischen 58 Minuten und 1 Stunde und 27 Minuten; die durchschnittliche maximale Bearbeitungszeit lag bei 1 Stunde und 9 Minuten.

Fehlende Werte (*missing data*) können eine statistische Auswertung der erhobenen Daten erschweren (Bortz & Döring, 2006, S. 236), wobei unterschiedliche Strategien zum Umgang mit fehlenden Werten diskutiert werden (für eine Übersicht vgl. Lüdtke, Robitzsch, Trautwein & Köller 2007). Im schriftlichen Teil der vorliegenden Studie werden fehlende Werte als ‚falsch‘ gewertet, wenn sie sich nicht auf Zeitprobleme zurückführen lassen, da die Probanden aufgefordert waren, alle Fragen zu beantworten und den Testbogen vollständig

ausgefüllt abzugeben. Die Bearbeitungszeit war in diesem Testbestandteil nicht begrenzt und die Probanden waren nicht zum Raten der richtigen Antwort angehalten, weshalb davon ausgegangen wird, dass nicht bearbeitete Fragen nicht aufgrund von Zeitdruck unbeantwortet geblieben sind, sondern dass die Probanden die richtige Antwort in diesen Fällen nicht wussten.

In der Fragebatterie zum deklarativen politikdidaktischen Wissen waren im Computerprogramm demgegenüber Zeitvorgaben zu jedem Item vorhanden. Die fehlenden Werte in dieser Fragebatterie kamen jedoch nicht durch Überschreiten der jeweiligen Bearbeitungszeit zustande. Die Antwortdauer liegt in allen Fällen bei null Sekunden, weshalb ein gesondertes Wertelabel für technisch bedingt fehlende Werte im Computerprogramm vKid definiert wurde (*90 = fehlender Wert [technisch oder $AW < 2$ Sek.]*). Technisch bedingt fehlende Werte liegen dann vor, wenn die Bearbeitungszeit des jeweiligen Items weniger als zwei Sekunden beträgt und daher davon ausgegangen werden kann, dass die Testpersonen das Item versehentlich übersprungen haben (Lindmeier, 2011, S. 161f.), oder aber, wenn technische Probleme (z. B. Fehler beim Abspielen einer Videovignette) von den Probanden berichtet bzw. in der Datenauswertung erkannt werden (bspw. fehlende Audioaufzeichnung einer Antwort der Probanden). Technisch bedingt fehlende Werte müssen im Gegensatz zu Items, die nicht beantwortet wurden, in der Auswertung wie nicht vorhandene Werte behandelt werden. In der Fragebatterie zum prozeduralen politikdidaktischen Wissen meldeten drei Probanden (ID_15/30; ID_20/29; ID_24/25) während des Tests beim Versuchsleiter solche Probleme an. Für diese Versuchspersonen wurde ein neuer Testdurchlauf unmittelbar nach dem ersten Durchlauf unter Ausschluss der bereits beantworteten Fragen angesetzt, um ihre Antworten dennoch in die Auswertung mit einbeziehen zu können. In der zeitlichen Auswertung zeigte sich jedoch, dass die Probanden in allen Fällen (insgesamt 14 Items) vor Auftreten der technischen Probleme das entsprechende Item mindestens vier Sekunden und maximal im kompletten Umfang der Bearbeitungszeit von drei Minuten angesehen hatten. Da sie somit im ersten Anlauf bspw. die zusätzlichen Hinweise zur Beantwortung der Items bereits gelesen haben könnten, dürfen Bearbeitungsvorteile für den zweiten Durchlauf nicht vollumfänglich ausgeschlossen werden, sodass diese Antworten dem Wertelabel *90 = fehlender Wert (technisch oder $AW < 2$ Sek.)* zugeordnet und damit aus dem Auswertungsverfahren ausgeschlossen wurden. Der Anteil technisch bedingt fehlender Werte im Computerprogramm liegt insgesamt bei etwas mehr als einem Prozent. Dabei kann

der höhere Anteil technisch bedingt fehlender Werte in der Fragebatterie zum prozeduralen politikdidaktischen Wissen durch die größere Fehleranfälligkeit computerbasierter Verfahren bei zunehmendem technischem Aufwand erklärt werden (vgl. Tabelle A. 10 und Tabelle A. 11 im Anhang für eine Übersicht über technisch bedingt fehlende Werte).

Im Gegensatz zu den Fragebatterien zum Fachwissen und zum deklarativen politikdidaktischen Wissen wird der Zeitdruck in der Skala zum prozeduralen politikdidaktischen Wissen als wichtiger qualitativer Einflussfaktor auf die *Spontaneität* und *Unmittelbarkeit* der Antworten der Probanden angesehen (vgl. auch Kapitel 3.2.2 und Kapitel 6.3.1). Sofern in dieser Fragebatterie keine Reaktion der Probanden erfolgte, die Antwortdauer jedoch über zwei Sekunden liegt und keine technischen Probleme erkennbar sind, wird dies als inadäquate Rückmeldung bewertet.

7.2.2 Urteilsübereinstimmung der Rater

Das auf Grundlage des Ratingmanuals durchgeführte Beurteilungsverfahren der Antworten auf die 10 Items zum prozeduralen politikdidaktischen Wissen basiert auf einer 0/1/2-Skala (Lindmeier, 2011; Rehm & Bölsterli, 2014, S. 220; vgl. auch Kapitel 6.3.2.1). Hierbei ist die Konkordanz, also das Übereinstimmungsmaß zwischen den Beurteilenden, ein zentraler Referenzpunkt dafür, ob die Urteile der Rater zu einem Gesamturteil zusammengefasst werden können (Bortz & Döring, 2006, S. 185); sie ist somit ein Gütekriterium für die Objektivität (Schmidt-Atzert & Amelang, 2012, S. 319). Zur Bestimmung der Konkordanz der Urteile bei kategorialen Daten und zwei Ratern wird in der Regel der Cohens Kappa-Koeffizient (κ) bestimmt (Bortz & Lienert, 2008, S. 310). Die so ermittelte Maßzahl bewegt sich für die Items zum prozeduralen politikdidaktischen Wissen zwischen $\kappa = .62$ und $\kappa = .79$ (vgl. Tabelle A. 12 im Anhang). Nach Bortz und Döring (2006, S. 277) sowie Landis und Koch (1977) lassen die Werte auf eine gute bzw. beachtliche Übereinstimmung zwischen den Ratern schließen.

Da sich die Beurteilungskategorien in einer Rangordnung darstellen lassen und die Beurteilungsskala somit auf ordinalem Niveau liegt, kann an der Berechnung von Cohens Kappa problematisiert werden, dass unter Umständen vorliegende größere Abweichungen (Diskordanzen) in der Übereinstimmung der Rater (z. B. wenn die Antwort eines Probanden von Rater A als *adäquat* und von Rater B als *inadäquat* beurteilt wird) nicht berücksichtigt wer-

den, da alle fehlenden Übereinstimmungen gleich ‚schwer‘ verrechnet werden. Daher wird die Berechnung des gewichteten Cohens Kappa (Weighted-Kappa-Koeffizient κ_w) vorgeschlagen (Bortz & Lienert, 2008, S. 321), bei dem anhand eines Gewichtungsfaktors Nichtübereinstimmungen differenziell gewichtet werden. Die Werte des linear gewichteten Cohens Kappa liegen für die Items zum prozeduralen politikdidaktischen Wissen zwischen $\kappa_w = .66$ und $\kappa_w = .79$ und des quadratisch gewichteten Cohens Kappa zwischen $\kappa_w = .66$ und $\kappa_w = .80$ (vgl. Tabelle A. 12 im Anhang).⁵⁰ Die gewichteten Kappa-Werte liegen relativ zu den entsprechenden ungewichteten Kappa-Werten nach Cohen – für Letztere mit der Ausnahme des Items FpW-I_2 – mindestens auf gleichem, mehrheitlich auf leicht höherem Niveau. Dies kann als Indikator dafür angesehen werden, dass gravierende Diskordanzen seltener vorkommen als weniger gravierende. Insgesamt bestätigen die gewichteten Cohens Kappa-Werte eine gute Übereinstimmung zwischen den Ratern in ihrer Beurteilung der Antworten auf Fragen zum prozeduralen politikdidaktischen Wissen.

Schließlich kann bei ordinalskalierten Daten und zwei Beurteilern als Maß der Urteilskonkordanz der Rangkorrelationskoeffizient Rho von Spearman (ρ) bestimmt werden (Bortz, Lienert & Boehnke, 2008, S. 465). Für die Berechnung wird davon ausgegangen, dass die Positionen der Ränge für die vorliegende Rangfolge äquidistant sind (ebd., S. 414).⁵¹ Das so berechnete Spearmans Rho weist durchgängig signifikante positive Korrelationen zwischen den Variablen aus, was ebenfalls auf eine hohe Urteilsübereinstimmung zwischen den Ratern schließen lässt (vgl. Tabelle A. 12 im Anhang).

7.2.3 Itemanalyse

Die genaue statistische Betrachtung von Items soll Aussagen über deren Qualität erlauben, da „[d]ie Itemanalyse (Aufgabenanalyse) ... ein zentrales Instrument der Testkonstruktion und Testbewertung [ist]“ (Bortz & Döring, 2006, S. 217). In der nachfolgenden Analyse

⁵⁰ Als relative Distanz zwischen den Stufen der 0/1/2-Skala (Ordinalniveau) für die Berechnung des gewichteten Cohens Kappa wird der Wert *eins* festgelegt (Bortz & Lienert, 2008, S. 321).

⁵¹ Da Spearmans „ ρ ... auf dem Konzept [beruht], die Rangskalen als Kardinalskalen aufzufassen und die Ränge rechnerisch wie Messwerte zu behandeln“ (Bortz et al., 2008, S. 414), ist diese Annahme in Bezug auf die durch die Rangwerte repräsentierten Merkmalsausprägungen nicht „trivial“ (ebd.), kann jedoch für die vorliegenden Daten begründet werden (vgl. Kapitel 6.3.2).

werden dabei die Schwierigkeiten der Items (Kapitel 7.2.3.1) sowie deren Trennschärfe (Kapitel 7.2.3.2) untersucht.

7.2.3.1 Itemschwierigkeit

Die Frage nach der Beantwortungsschwierigkeit der einzelnen Items wird durch Itemschwierigkeitsindizes (p_i) ausgedrückt (ebd., S. 218), wobei p_i Werte zwischen null und eins annehmen kann. Je höher der Wert, desto mehr Probanden haben das Item richtig beantwortet (der Schwierigkeitsindex $p_i = 1$ würde aussagen, dass alle Testpersonen das jeweilige Item richtig beantwortet haben) (Schmidt-Atzert & Amelang, 2012, S. 114). Die Items zum Fachwissen sowie zum deklarativen politikdidaktischen Wissen wurden dichotom kodiert, sodass die Itemschwierigkeit direkt berechnet werden kann.⁵²

Zur Frage, ob bei der Berechnung der Schwierigkeitsindizes von Multiple-Choice-Items zu berücksichtigen ist, dass Antworten auch erraten oder generell zufällig richtig oder falsch sein könnten (Rate- bzw. Zufallskorrektur), gibt es in der Literatur unterschiedliche Auffassungen. So kommen Schmidt-Atzert und Amelang zu dem Schluss, dass „[d]ie Ratekorrektur ... nur dann zur Anwendung kommen [sollte], wenn in der Instruktion zum Raten aufgefordert wird für den Fall, dass man die Lösung nicht findet“ (ebd., S. 116). Die Autoren erwarten, dass sogar für Antworten mit lediglich zwei Antwortalternativen (50 % Ratewahrscheinlichkeit) keine Ratekorrektur durchgeführt wird (ebd., S. 117). Bortz und Döring (2006) betonen demgegenüber die Notwendigkeit einer Ratekorrektur (S. 216f.), wobei sie darauf hinweisen, dass „[d]ie Bedeutsamkeit des Ratens für die Reliabilität ... von der Testlänge, der Trefferwahrscheinlichkeit und der Personenquote, die zum Raten aufgefordert wurde, ab[hängt]“ (ebd., S. 217). Bühner (2006) empfiehlt, „Raten schon durch die Auswahl geeigneter Antwortformate und geeigneter Distraktoren zu unterbinden, so dass keine Ratekorrektur mehr nötig ist“ (S. 85). Auch Hussy, Schreier und Echterhoff (2013) gehen davon aus, dass „[b]ei Multiple-Choice-Items ... die Wahrscheinlichkeit solcher Verfälschungen verringert werden [kann], indem neben der korrekten Antwort geeignete Distraktoren eingesetzt werden“ (S. 88; vgl. auch Kapitel 6.3). Eben solche Distraktoren sollten eine ausreichende Attraktivität sowie eine hohe Nachvollziehbarkeit aufweisen (vgl. auch Moosbrugger &

⁵² $p_i = \frac{R_i}{N_i}$, wobei R für die Zahl der Probanden steht, die das Item i korrekt beantwortet haben und N_i für die Zahl aller Probanden, die das Item beantwortet haben (Bortz & Döring, 2006, S. 218).

Kelava, 2012, S. 45). Probanden, die die richtige Antwort nicht wissen, sind dann „in der Regel davon *überzeugt*, die richtige Antwort gewählt zu haben und müssen daher gar nicht erst raten bzw. tun dies nur in wenigen Fällen“ (Kemna, 2012, S. 187). Lienert und Raatz (1998) schlussfolgern nach einer Zusammenschau von Ergebnissen unterschiedlicher Studien zu den Auswirkungen des Ratens auf die Testreliabilität, „daß [*sic*] man keinesfalls zum blinden Raten auffordern sollte, sondern ein gewisses ‚Fingerspitzengefühl‘ beim Suchen der Lösung empfehlen sollte, und daß [*sic*] man dann lieber auf die Zufallskorrektur verzichten sollte“ (S. 47). Sie resümieren, dass eine Korrektur nur dann durchgeführt werden sollte, wenn in den Testanweisungen explizit zum Raten aufgefordert wird, insgesamt nur wenige Aufgaben bearbeitet werden oder die Rohtestwerte Hinweise auf häufiges Raten enthalten (ebd., S. 69). Da für die vorliegende Studie in präpilotierenden Expertensitzungen die Qualität der Distraktoren optimiert wurden, die Fragebatterien einen großen Umfang haben und zudem die Rohtestwerte nicht erkennen ließen, dass einzelne Probanden wahllos Antworten angekreuzt haben, wird nicht davon ausgegangen, dass eine Ratekorrektur der Werte notwendig ist.

Die Berechnung der Itemschwierigkeiten zeigt schließlich, dass sich etwas weniger als zwei Drittel aller erhobenen Items zum Fachwissen in einem mittleren Bereich zwischen $p_i = .20$ und $p_i = .80$ (Bortz & Döring, 2006, S. 219) bewegen; 15 Items liegen außerhalb dieses Wertebereichs (vgl. Tabelle A. 13 im Anhang). Insgesamt drei Fünftel aller erhobenen Items zum deklarativen politikdidaktischen Wissen weisen ein mittleres Schwierigkeitsniveau auf; acht Items zu dieser Fragebatterie liegen außerhalb eines mittleren Schwierigkeitsbereichs (Tabelle A. 14 im Anhang). Für die Items zum prozeduralen politikdidaktischen Wissen liegen die Beurteilungen der Rater in Form einer mehrstufigen 0/1/2-Skala vor. Für mehrstufige Items geben Bortz & Döring (2006, S. 219) eine Berechnungsformel an, bei der sich der Schwierigkeitsindex eines Items als Summe aus den erreichten Punkten aller Probanden geteilt durch die maximal mögliche Punktesumme als Produkt der maximalen Punktzahl aller antwortenden Personen ergibt.⁵³ Alle Items zum prozeduralen politikdidaktischen Wissen liegen demnach in einem mittleren Schwierigkeitsbereich (vgl. Tabelle A. 15 im Anhang).

⁵³ $p_i = \frac{\sum_{m=1}^n x_{im}}{k_i \cdot n}$, wobei x_i für die erreichten Punkte auf Item i steht und k für die maximale Punktzahl (Bortz & Döring, 2006, S. 219).

Im Hinblick auf die in den Skalen zum Fachwissen sowie zum deklarativen fachdidaktischen Wissen außerhalb eines mittleren Schwierigkeitsbereiches liegenden Items ist zu berücksichtigen, dass Items von den Skalenendpunkten (extrem schwierig bzw. extrem leicht) nur geringe Aussagekraft haben, da sie zur Differenzierung zwischen Personen nicht geeignet sind (Rentzsch & Schütz, 2009, S. 201). Schmidt-Atzert und Amelang (2012) weisen jedoch darauf hin, dass, „[s]oll ein Test auch im unteren und oberen Merkmalsbereich gut zwischen den Personen differenzieren, ... man auch Items [benötigt], die für ‚normale‘ Personen schwer bzw. leicht sind“ (S. 114), da „Items, die bei Personen mit mittlerer Merkmalsausprägung optimal differenzieren, ... bei Personen mit extremer Merkmalsausprägung eine sehr geringe Streuung [haben]“⁵⁴ (ebd.), womit auch die Abhängigkeit der Itemschwierigkeit von der Stichprobe angesprochen ist (ebd., S. 119; für eine kritische Reflexion einer solchen Stichprobenabhängigkeit vgl. Kapitel 7.3).

7.2.3.2 Itemtrennschärfe

Anhand der Ermittlung des Trennschärfekoeffizienten (r_{it}) lässt sich erkennen, „wie gut ein einzelnes Item das Gesamtergebnis repräsentiert“ (ebd.). Es zeigt zudem an, „wie gut das Item zwischen Personen mit hoher versus niedriger Merkmalsausprägung ‚trennt‘“ (Schmidt-Atzert & Amelang, 2012, S. 119). Der Trennschärfekoeffizient kann Werte zwischen -1 und +1 annehmen, wobei Bortz und Döring (2006, S. 220) für mittelmäßige Werte den Bereich zwischen $r_{it} = .30$ und $r_{it} = .50$ angeben und für hohe Werte den Bereich $r_{it} > .50$. Dabei hängt die Trennschärfe der Items mit deren Schwierigkeiten und der Itemstreuung, aber auch mit der Streuung der Testwerte ab. Falls „der Test einer sehr homogenen Stichprobe von Personen vorgegeben [wurde], bei der das Merkmal nicht stark variiert, müssen die Trennschärfen zwangsläufig niedriger ausfallen als bei einer repräsentativen Stichprobe“ (Schmidt-Atzert & Amelang, 2012, S. 122).

Die Trennschärfen für die Items zum Fachwissen liegen zwischen $r_{it} = -.27$ und $r_{it} = .70$, für die Items zum deklarativen politikdidaktischen Wissen zwischen $r_{it} = -.14$ und $r_{it} = .46$ und für die Items zum prozeduralen politikdidaktischen Wissen zwischen $r_{it} = .08$ und $r_{it} = .39$.

⁵⁴ Die Itemstreuung bzw. Standardabweichung eines Items (s_i)⁵⁴ ist bei dichotomen Werten rechnerisch durch die Itemschwierigkeit bestimmt (Bühner, 2006, S. 86), sie „ist also bei einer mittleren Schwierigkeit ($p = .5$) am größten und in der Nähe von null und eins sehr niedrig“ (Schmidt-Atzert & Amelang, 2012, S. 114).

Diese Befunde – vgl. auch Tabelle A. 13, Tabelle A. 14 und Tabelle A. 15 im Anhang für eine detaillierte Übersicht über die Trennschärfen aller Test-Items – sind unter Berücksichtigung der angesprochenen Abhängigkeit der Trennschärfekoeffizienten von der Itemschwierigkeit sowie Itemvarianz insofern nachvollziehbar, als die Analyse der Schwierigkeiten der Items zum Fachwissen und zum deklarativen politikdidaktischen Wissen insgesamt 23 Items außerhalb eines mittleren Schwierigkeitsbereiches ausgewiesen hat.

Grundsätzlich sind Items nicht brauchbar, die eine im niedrigen positiven Wertebereich oder im negativen Wertebereich liegende Trennschärfe aufweisen (Bortz & Döring, 2006, S. 220). Jedoch, so geben Schmidt-Atzert und Amelang (2012, S. 123) zu bedenken, sollten Items nicht unüberlegt wegen niedriger Trennschärfe aus der Itemanalyse ausgeschlossen werden. Sie führen dazu einen möglichen Verlust an Validität an, der mit einer Erhöhung der Reliabilität durch das Aussondern von Items mit niedrigem Trennschärfekoeffizienten einhergehen könne (ebd.). So würde der Ausschluss bei einer Orientierung an hoch korrelierenden Items zwar für eine Homogenisierung des Tests und einer Steigerung der internen Konsistenz sorgen, es sei aber stets zu prüfen, ob die verbliebenen Items ein Merkmal dann noch in seiner konzeptualisierten inhaltlichen Breite erfassen oder nur noch in einem begrenzten Ausschnitt (ebd. S. 123f.). Sofern zudem die Leistungen in den Randbereichen interessieren, fügt Bühner (2006) an: „Führt man eine Itemselektion nur anhand der Itemtrennschärfen durch, besteht die Gefahr, in hohem Maße Items mit extremer Schwierigkeit aus dem Test zu eliminieren“ (S. 99). Es wird daher vorgeschlagen, die Trennschärfe von Test-Items im Zusammenhang mit ihrer Schwierigkeit zu analysieren (ebd., S. 106f.; Schmidt-Atzert & Amelang, 2012, S. 122). Weil Differenzierungsmöglichkeiten in den extremen Schwierigkeitsbereichen für die vorliegende Arbeit von Interesse sind, soll eine solche Zusammenhangsanalyse dabei helfen, die für die nachfolgenden Betrachtungen besonders relevanten Items zu identifizieren. Im Folgenden ist dies beispielhaft anhand der Skala zum Fachwissen dargestellt (Abbildung 24).

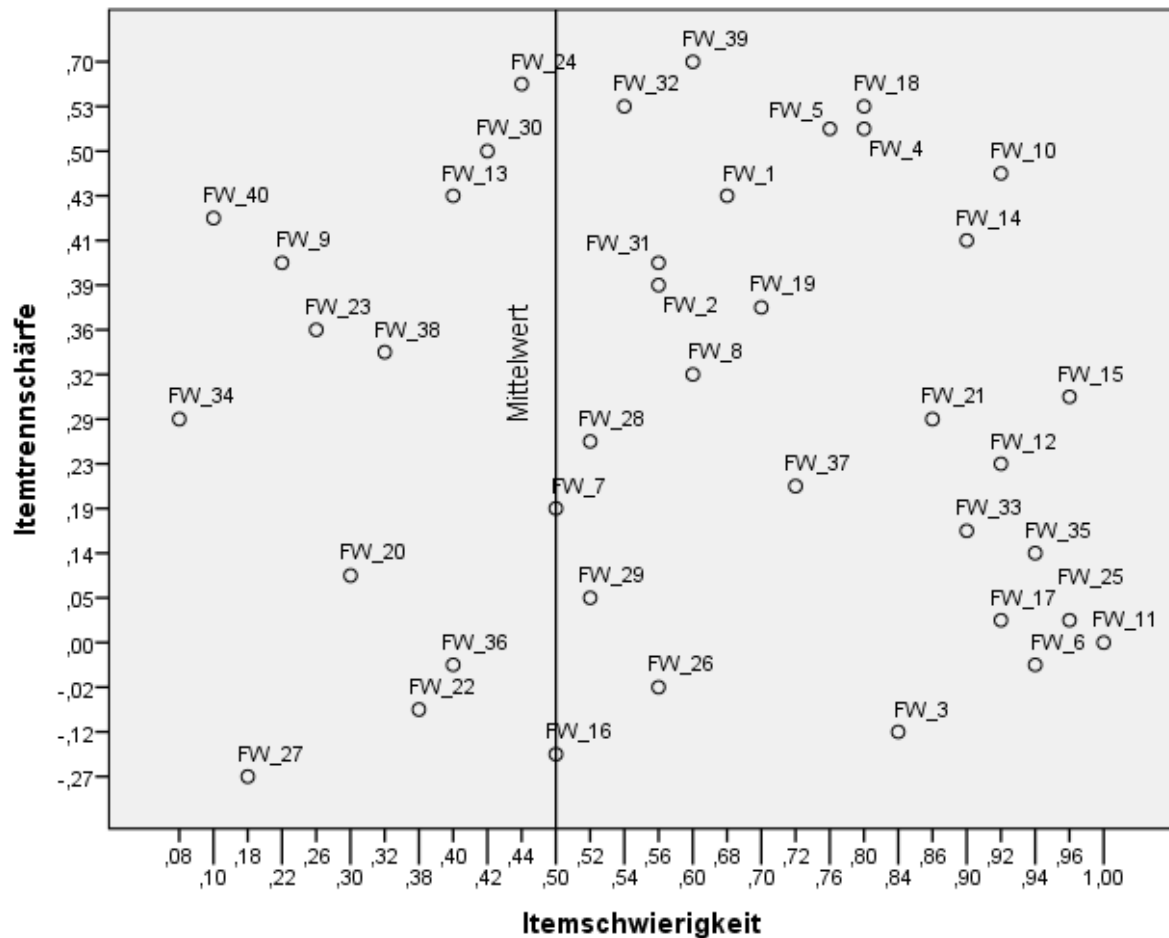


Abbildung 24: Zusammenhang zwischen Itemschwierigkeit und Trennschärfe in den vorliegenden Daten zum Fachwissen (FW)

Es zeigt sich, dass einige Items hohe bzw. niedrige Schwierigkeiten aufweisen, da sie an den Endpunkten der Skala liegen. Gleichzeitig jedoch besitzen einige Items trotz gleicher Distanz vom theoretischen Mittelpunkt der dichotomen Skala sehr unterschiedliche Trennschärfen. So liegen die Items FW_10, FW_12 und FW_17 gleich weit rechts vom theoretischen Mittelpunkt der Skala entfernt und können damit offenbar als leichte Items identifiziert werden, allerdings weist das Item FW_10 eine erheblich höhere Trennschärfe auf als die beiden anderen Items. Ähnliches lässt für die Items FW_27 und FW_40 feststellen, bei denen die Trennschärfekoeffizienten trotz etwa gleicher Distanz links vom theoretischen Mittelpunkt der Skala aus ebenfalls wesentlich voneinander abweichen. Beide Items können als schwierig eingestuft werden; während jedoch das Item FW_27 negativ mit der Skala korreliert und daher augenscheinlich etwas anderes misst als die Skala, weist das Item FW_40 eine höhere

Trennschärfe auf und misst daher mit höherer Wahrscheinlichkeit das Gleiche wie die Skala bzw. differenziert besser zwischen Probanden mit hoher und niedriger Merkmalsausprägung (Schmidt-Atzert & Amelang, 2012, S. 119). Es scheint daher angezeigt, zur Differenzierung in den Randbereichen der Skala nicht alle Items zum Fachwissen mit einem Schwierigkeitskoeffizienten $p_i = 5$ bis $p_i = 10$ bzw. $p_i = 90$ bis $p_i = 95$ aus der weiteren Item- und Skalenanalyse auszuschließen, sofern diese nicht zu Einbußen an Trennschärfe führen ($r_{it} \geq .30$) (Amelang & Schmidt-Atzert, 2006, S. 119f.). Gleichzeitig ist zu berücksichtigen, dass ein Item nur dann aus der Skala entfernt werden sollte, wenn dies mit inhaltlich fehlender Passgenauigkeit gerechtfertigt werden kann (Bühner, 2006, S. 106f.). Abbildung 25 zeigt den Zusammenhang von Itemtrennschärfe und Itemschwierigkeit in der Fragebatterie zum Fachwissen, in der schließlich schlecht passende Items ausgeschlossen wurden.

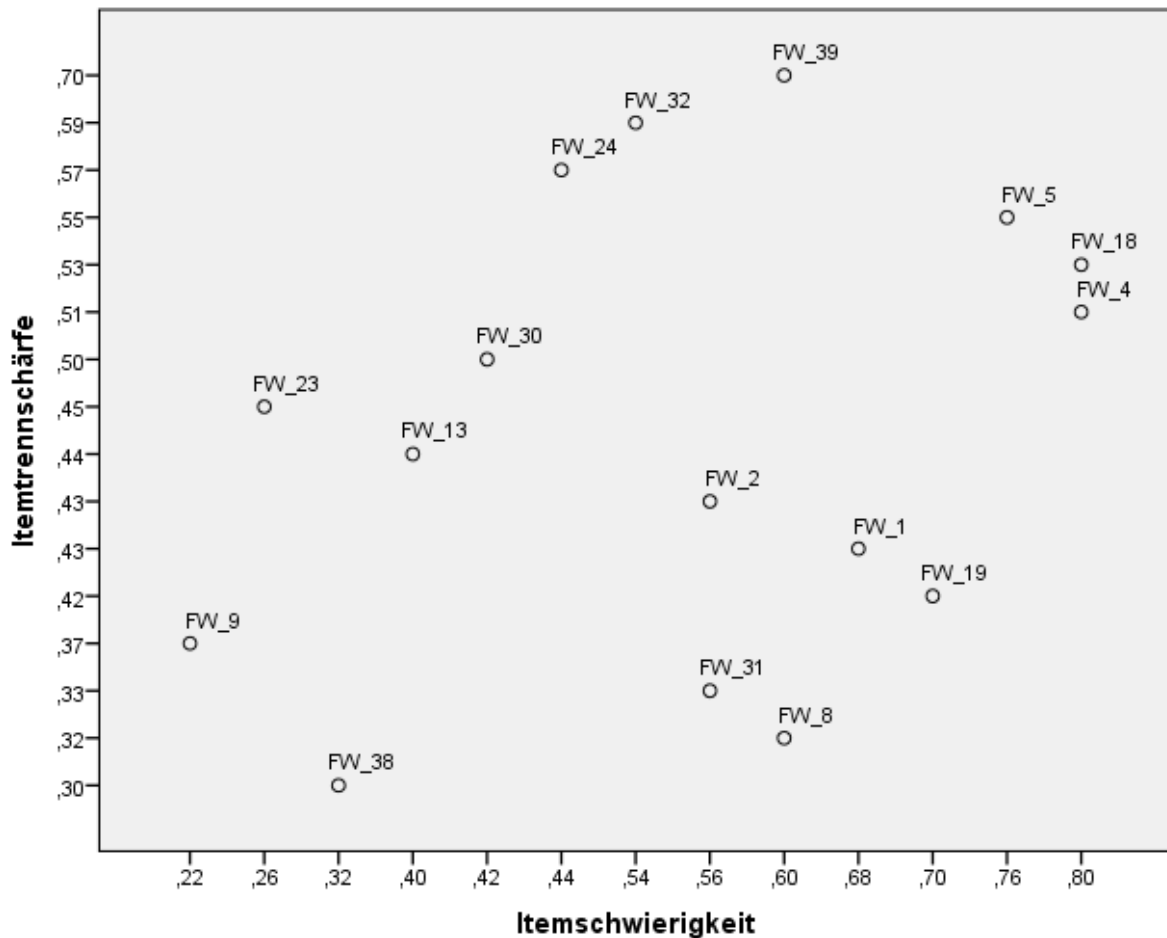


Abbildung 25: Zusammenhang zwischen Itemschwierigkeit und Trennschärfe in den vorliegenden Daten zum Fachwissen unter Berücksichtigung ausgeschlossener Items (FW_rev)

Anmerkung: Die Itemtrennschärfe weist die korrigierte Item-Skala-Korrelation unter Berücksichtigung der ausgeschlossenen Items aus; es wurden 24 Items ausgeschlossen.

Hinsichtlich der Trennschärfe der Test-Items in den drei Fragebatterien lässt sich zusammenfassen, dass diese nicht durchgehend in einem mittleren Bereich liegt. Dabei konnten die Zusammenhänge zwischen Itemtrennschärfe und Itemschwierigkeit am Beispiel der Items in der Skala zum Fachwissen aufgezeigt und deren teilweise zu identifizierenden niedrigen Ausprägungsgrade somit vorläufig erklärt werden. Gleichzeitig wurde begründet, warum der Ausschluss von Items in den folgenden Item- und Skalenanalysen ausschließlich auf Grundlage niedriger Itemtrennschärfen nicht gerechtfertigt ist, da eine Erhöhung der Reliabilität einer Testskala durch Ausschluss von weniger trennscharfen Items ggf. einen Verlust an In-

haltsvalidität nach sich ziehen kann, wenn die Itemauswahl nicht inhaltlich begründet ist (Schmidt-Atzert & Amelang, 2012, S. 123). Das angeführte Vorgehen ist daher in den weiteren Analyseschritten für die Items zum deklarativen und prozeduralen politikdidaktischen Wissen ebenfalls zu berücksichtigen; vgl. auch Abbildung A. 2 und Abbildung A. 3 im Anhang für eine Darstellung des Zusammenhangs zwischen Itemtrennschärfe und Itemschwierigkeit in den Skalen zum deklarativen politikdidaktischen Wissen und zum prozeduralen politikdidaktischen Wissen.

7.2.4 Interne Konsistenz der Skalen

Um die Reliabilität der Skalen bestimmen zu können, wird als häufigstes Maß der internen Konsistenz der Alphakoeffizient (α) von Cronbach bestimmt, wobei Alpha „der mittleren Testhalbierungsreliabilität eines Tests für alle möglichen Testhalbierungen [entspricht]“ (Bortz & Döring, 2006, S. 198). Zwar „wird dieses Maß zuweilen auch als Homogenitätsindex verwendet“ (ebd.), da Alpha mit zunehmender Item-Anzahl größer wird, jedoch weisen Schmidt-Atzert und Amelang (2012) darauf hin, dass „[d]araus ... *nicht* im Umkehrschluss [folgt], dass ein hohes Alpha für eine große Homogenität des Tests spricht“ (S. 49), weil zum Beispiel „ein Test für ein heterogenes Konstrukt, der aus niedrig korrelierenden Items besteht, ... deshalb bei sehr vielen Items ein hohes Alpha aufweisen [kann]“ (ebd., S. 50). Es gelte weiterhin zu berücksichtigen, dass ein hohes Cronbachs Alpha nicht grundsätzlich anzustreben sei, wenn bspw. eine Skala eher heterogene Merkmale erfassen soll (ebd., S. 49).

Es werden unterschiedliche Anforderungen genannt, wie hoch ein guter Alpha-Wert sein sollte. Bortz und Döring (2006, S. 199) geben als unteren Schwellenwert einen Alphakoeffizienten von $\alpha = .80$ sowie für hohe Reliabilitäten einen Alphakoeffizienten von $\alpha > .90$ an, wobei diese Werte nicht für explorative Zwecke gelten würden. Schmitt (1996) beschreibt einen Alphakoeffizienten von $\alpha = .70$ als anstrebenswert, betont jedoch unter anderem wegen der Abhängigkeit der Alpha-Werte von der Skalenlänge gleichzeitig: „the use of any cutoff value (including .70) is shortsighted“ (S. 351). Teilweise werden für explorative Studien akzeptable untere Alpha-Schwellenwerte von $\alpha = .60$ bis $\alpha = .70$ angegeben (Hair, Black, Babin & Anderson, 2010, S. 125).

Zur Reliabilitätsbestimmung bei mehrdimensionalen Tests empfehlen Bortz und Döring (2006, S. 199), die interne Konsistenz der Skalen separat zu bestimmen, statt einen gemein-

samen Reliabilitätskoeffizienten für den Gesamttest zu berechnen. Um hierbei die auf unterschiedlicher Skalenlänge basierenden Alpha-Werte besser miteinander vergleichen zu können, extrapoliert Lindmeier (2011, S. 166) in ihrer Darstellung der internen Konsistenz der Skalen zum mathematischen Basiswissen, zur Reflexionskompetenz sowie zu den aktionsbezogenen Kompetenzen fehlende Items mithilfe der Spearman-Brown-Prophecy-Formel⁵⁵ zur Aufwertung des Reliabilitätskoeffizienten. Nach Bortz und Döring (2006) kann diese Formel einen „nach der Testhalbierungsmethode gewonnene Reliabilitätskoeffizient ... nachträglich um den Betrag, der durch die Testhalbierung verloren ging, [aufwerten]“ (S. 198). Da sich die Spearman-Brown-Prophecy-Formel demnach auf die Testhalbierungsreliabilität bezieht, kann der nach dieser Formel ermittelte Reliabilitätskoeffizient für die vorliegenden Daten lediglich als grobe Schätzung angesehen werden (Lindmeier, 2011, S. 166). Er soll im Folgenden jedoch zwecks besserer Vergleichbarkeit der Alpha-Werte ebenfalls als Schätzwert ermittelt werden (Tabelle 3).

Tabelle 3: Interne Konsistenz der (revidierten) Skalen sowie des Gesamttests (Reliabilitätskoeffizient α , Standardmessfehler SE und auf 16 Items geschätzter Alpha-Wert)

Skala	Itemanzahl	α (Cronbach)	SE	α geschätzt
FW	40	.78	.13	-
FdW	20	.51	.16	-
FpW	10	.58	.04	-
FW_rev	16	.85	.07	-
FdW_rev	10	.63	.14	.73
FpW_rev	8	.61	.04	.76
Alle	70	.70	.14	-
Alle_rev	34	.70	.12	.52
FW_rev UND FdW_rev	26	.85	.08	.78
FW_rev UND FpW_rev	24	.66	.10	.56
FdW_rev UND FpW_rev	18	.20	.21	.18

⁵⁵ $Rel_{\text{kor}} = \frac{k \cdot Rel}{1 + (k-1) \cdot Rel}$, wobei Rel die Reliabilität des Tests ist (α) und k für den Wert steht, um den sich die Itemanzahl erhöht (Schmidt-Atzert & Amelang, 2012, S. 48f.).

Die Aufstellung zeigt, dass Cronbachs Alpha für die Skala zum *Fachwissen* mit allen Items (FW) mit $\alpha = .78$ in einem akzeptablen, tendenziell guten Bereich liegt. Der Ausschluss schlecht passender Items erhöht den Reliabilitätskoeffizienten für die Skala (FW_rev) auf $\alpha = .85$; dieser Wert kann als gut angesehen werden. Die Skala zum *deklarativen politikdidaktischen Wissen* ohne Ausschluss von Items (FdW) weist einen Reliabilitätswert von $\alpha = .51$ auf. Dieser Wert muss – auch wegen des relativ großen Itemumfangs der Skala – als fragwürdig angesehen werden. Der Ausschluss schlecht passender Items erhöht den Reliabilitätskoeffizienten für diese Skala (FdW_rev) auf $\alpha = .63$, was in Anbetracht der relativ geringen Skalengröße sowie des explorativen Zwecks als noch akzeptabler Wert angesehen werden kann. Wird probeweise mithilfe der Spearman-Brown-Prophecy-Formel der auf einen höheren Itemumfang geschätzte Alpha-Wert berechnet, ergibt sich ein Reliabilitätskoeffizient von $\alpha = .73$. Die Skala zum *prozeduralen politikdidaktischen Wissen* mit allen Items (FpW) weist ein Cronbachs Alpha von $\alpha = .58$ auf. Die Itemanalyse zeigt für diese Skala im Vergleich zu den beiden anderen Skalen einen insgesamt höheren durchschnittlichen Schwierigkeitsindex aller Items verbunden mit einer relativ niedrigen Varianz der Antworten. Aus dieser Skala wurden zwei Items mit sehr niedriger Trennschärfe entfernt. Um einen gewissen Mindestumfang der Skala beizubehalten, wurden relativ geringe Trennschärfen der Items toleriert (vgl. auch Abbildung A. 3 im Anhang). Unter Ausschluss unpassender Items zeigt die revidierte Skala zum prozeduralen politikdidaktischen Wissen (FpW_rev) ein insgesamt noch akzeptables Cronbachs Alpha von $\alpha = .61$ sowie nach Extrapolation dieses Wertes einen Reliabilitätskoeffizienten von $\alpha = .76$.

Wenn probeweise Cronbachs Alpha für alle drei Skalen (70 Items) berechnet wird, zeigt sich eine niedrigere Reliabilität der Gesamtskala ($\alpha = .70$) im Vergleich zur Reliabilität der Subskala FW sowie – unter Berücksichtigung, dass der Reliabilitätskoeffizient mit der Anzahl der Items grundsätzlich größer wird – eine lediglich leicht erhöhte Reliabilität im Vergleich zu den Subskalen FdW und FpW. Da sich Cronbachs Alpha für alle drei revidierten Skalen (Alle_rev) auch unter Berücksichtigung der fehlenden Werte nicht erhöht ($\alpha = .70$) und hier der Unterschied zu den Reliabilitätskoeffizienten aller revidierten Subskalen ebenfalls niedriger (FW_rev) bzw. nur marginal höher ist (FdW_rev; FpW_rev), kann dies tendenziell als Anzeichen dafür angesehen werden kann, dass die Gesamtskala entlang der Erwartungen an die Testkonstruktion auf Grundlage der Modellierung des politikdidaktischen Wissen in dieser Studie kein generelles Konstrukt misst (Lindmeier, 2011, S. 168). Noch deutlicher wird

dies, wenn Cronbachs Alpha für alle Skalenpaare berechnet wird. Eine solche Aufstellung lässt zwar ebenfalls lediglich vorläufige Aussagen zu, jedoch sinkt die interne Konsistenz der paarweise gebildeten revidierten Skalen teilweise deutlich in der Verbindung mit der Skala zum prozeduralen politikdidaktischen Wissen. Ebenso ist zu erkennen, dass die interne Konsistenz der gepaarten Skalen zum Fachwissen und zum deklarativen politikdidaktischen Wissen im Vergleich zu FW_rev nicht sinkt und im Vergleich zu FdW_rev ansteigt (vgl. Kapitel 7.3 zur Diskussion dieser Befunde).

Zusammenfassend lässt sich zur internen Konsistenz der Skalen festhalten, dass die Reliabilitätskoeffizienten nach Ausschluss unpassender Items für die revidierte Skala zum Fachwissen gute Werte und für die revidierten Skalen zum deklarativen politikdidaktischen Wissen sowie zum prozeduralen politikdidaktischen Wissen unter Berücksichtigung der explorativen Zwecke noch akzeptable Werte aufweisen. Die Berechnung von Cronbachs Alpha für die zu einem Gesamttest summierten Skalen lässt bis auf Weiteres die Aussage zu, dass dieser erwartungsgemäß offenbar kein generelles Konstrukt erfasst (ebd.). Eine paarweise Berechnung mit der Skala FpW_rev lässt den so ermittelten Reliabilitätswert trotz insgesamt größerer Itemanzahl sinken, was diese Annahme für diese Skala ansatzweise unterstützt.

7.2.5 Testschwierigkeit und Testgesamtleistungen

Um Aussagen zu den Testleistungen der Probanden treffen zu können, erscheint die Angabe der erreichten absoluten Gesamtpunkte in den einzelnen Skalen und im Gesamttest weniger geeignet, da es keine externen Bezugspunkte zur Beurteilung gibt. Stattdessen ist die Angabe der durchschnittlichen Testleistungen besser zur Betrachtung der erreichten Leistungsniveaus in den Skalen sowie im Gesamttest geeignet. Tabelle 4 zeigt die durchschnittlichen Testleistungen in jeder Skala sowie im Gesamttest; die Berechnung der Mittelwerte erfolgt von der *durchschnittlichen Testleistung je Item* aus. Die durchschnittlichen Testleistungen aller Probanden je Item in der Skala zum prozeduralen politikdidaktischen Wissen wurden zur der besseren Vergleichbarkeit mit den anderen Werten anhand der Formel zur Berechnung des Schwierigkeitsindex bei polytomen Skalen (Bortz & Döring, 2006, S. 219) transformiert. Innerhalb des so errechneten Wertebereiches liegt die höchste Leistung dann vor, wenn der

Wert *eins* erreicht wird (Lindmeier, 2011, S. 171).⁵⁶ Eine alternative Betrachtungsweise ist möglich, wenn die Berechnung der durchschnittlichen Testleistungen *je Proband über alle Items hinweg* vorgenommen wird und insofern die Personenfähigkeit betrachtet wird (Tabelle 5).

Tabelle 4: Durchschnittliche Testleistungen in den Skalen sowie im Gesamttest (kumulierte Testleistungen)

Skala	<i>N</i>	Itemanzahl	Min	Max	<i>M</i>	<i>SD</i>
FW	50	40	.08	1.00	.61	.27
FdW	50	20	.22	.96	.63	.23
FpW	50	10	.23	.43	.34	.06
Gesamt	50	70	.08	1.00	.58	.25
FW_rev	50	16	.22	.80	.54	.18
FdW_rev	50	10	.22	.96	.65	.23
FpW_rev	50	8	.23	.43	.34	.07
Gesamt_rev	50	34	.22	.96	.52	.21

Tabelle 5: Durchschnittliche Testleistungen in den Skalen sowie im Gesamttest (Personenfähigkeit)

Skala	<i>N</i>	Itemanzahl	Min	Max	<i>M</i>	<i>SD</i>
FW	50	40	.35	.93	.61	.14
FdW	50	20	.30	.85	.63	.13
FpW	50	10	.00	.65	.34	.13
Gesamt	50	70	.00	.93	.53	.19
FW_rev	50	16	.13	1.00	.54	.26
FdW_rev	50	10	.20	1.00	.65	.20
FpW_rev	50	8	.00	.75	.33	.15
Gesamt_rev	50	34	.00	1.00	.51	.25

⁵⁶ Hierbei wird davon ausgegangen, dass sich die Skalenwerte als Lösungsquote der Items interpretieren lassen (2 = hohe Lösungsquote; 1 = mittlere Lösungsquote; 0 = niedrige Lösungsquote) (Lindmeier, 2011, S. 164).

Die durchschnittliche Testleistung aller Probanden liegt in der Skala zum Fachwissen bei $M = .61$ ($SD = .27$), in der Skala zum deklarativen politikdidaktischen Wissen mit $M = .63$ ($SD = .23$) etwas höher und wesentlich niedriger in der Skala zum prozeduralen politikdidaktischen Wissen ($M = .34$; $SD = .06$). Für die revidierten Skalen unter Berücksichtigung der fehlenden Werte liegen die entsprechenden durchschnittlichen Testleistungen in der Fragebatterie FW_rev im Vergleich zur ursprünglichen Skala auf etwas niedrigerem Niveau ($M = .54$; $SD = .18$), auf lediglich leicht höherem Niveau in der Skala FdW_rev ($M = .65$; $SD = .23$) und auf gleichem, relativ niedrigem Niveau in der Skala FpW_rev ($M = .34$; $SD = .07$). Die Testschwierigkeit der Skala zum prozeduralen politikdidaktischen Wissen ist daher insgesamt höher. Die Verteilung der durchschnittlichen Testleistungen der Probanden (Personenfähigkeit) in den (revidierten) Skalen zeigen Abbildung A. 4 bis Abbildung A. 9 im Anhang. Alle durchschnittlichen Testleistungen der Probanden (Personenfähigkeit) sind innerhalb der erreichten Testwerte in den Skalen dabei normalverteilt. Zur Prüfung der Normalverteilungen wurde der Kolmogorov-Smirnov-Test (KS-Test) verwendet (Bortz et al., 2008, S. 319ff.). Die Werte der extremsten Differenzen der Skalen FW ($D = .126$; $p = .41$), FW_rev ($D = .136$; $p = .31$), FD ($D = .142$; $p = .27$), FdW_rev ($D = .118$; $p = .49$), FpW ($D = .137$; $p = .31$) sowie FpW_rev ($D = .145$; $p = .24$) liegen alle unter dem kritischen Wert⁵⁷ der Prüfgröße $D = .192$; zur grafischen Darstellung der Normalverteilung vgl. auch Abbildung A. 4 bis Abbildung A. 9 im Anhang.

Es kann vermutet werden, dass in der Skala zum prozeduralen politikdidaktischen Wissen höhere Testleistungen mit einer längeren Bearbeitungszeit einhergehen, „as giving a correct answer may take more time than e.g. giving no answer“ (Lindmeier, 2011, S. 172), wobei eine vollständige Korrelation zwischen Bearbeitungszeit und Testleistung anzeigen würde, dass eine längere Antwort grundsätzlich besser beurteilt würde als eine kürzere, was sich negativ auf die Inhaltsvalidität des Tests auswirken könnte (ebd.). Die Histogramme der Bearbeitungszeit für die (revidierte) Skala zum prozeduralen politikdidaktischen Wissen deuten auf eine Normalverteilung der Daten hin (Abbildung A. 10 und Abbildung A. 11 im Anhang); der KS-Test bestätigt die angenommene Normalverteilung sowohl für die Bearbeitungszeit ($D = .111$; $p = .57$) der Skala mit allen Items (FpW) als auch für die Bearbeitungszeit ($D = .085$; $p = .86$) der revidierten Skala (FpW_rev). Hierbei zeigen die Histogramme zu

⁵⁷ Die Approximation des Wertes für die Stichprobengröße $N = 50$ bei einer 5 %igen Irrtumswahrscheinlichkeit erfolgt anhand der Formel $\frac{1.36}{\sqrt{N}}$ (Bortz et al., 2008, S. 747).

den kumulierten Bearbeitungszeiten jeweils einen Ausreißer im unteren Bereich der Zeitskala, der auch als Ausreißer in den Streudiagrammen zur bivariaten Normalverteilung erkennbar ist (Abbildung 26 und Abbildung 27).

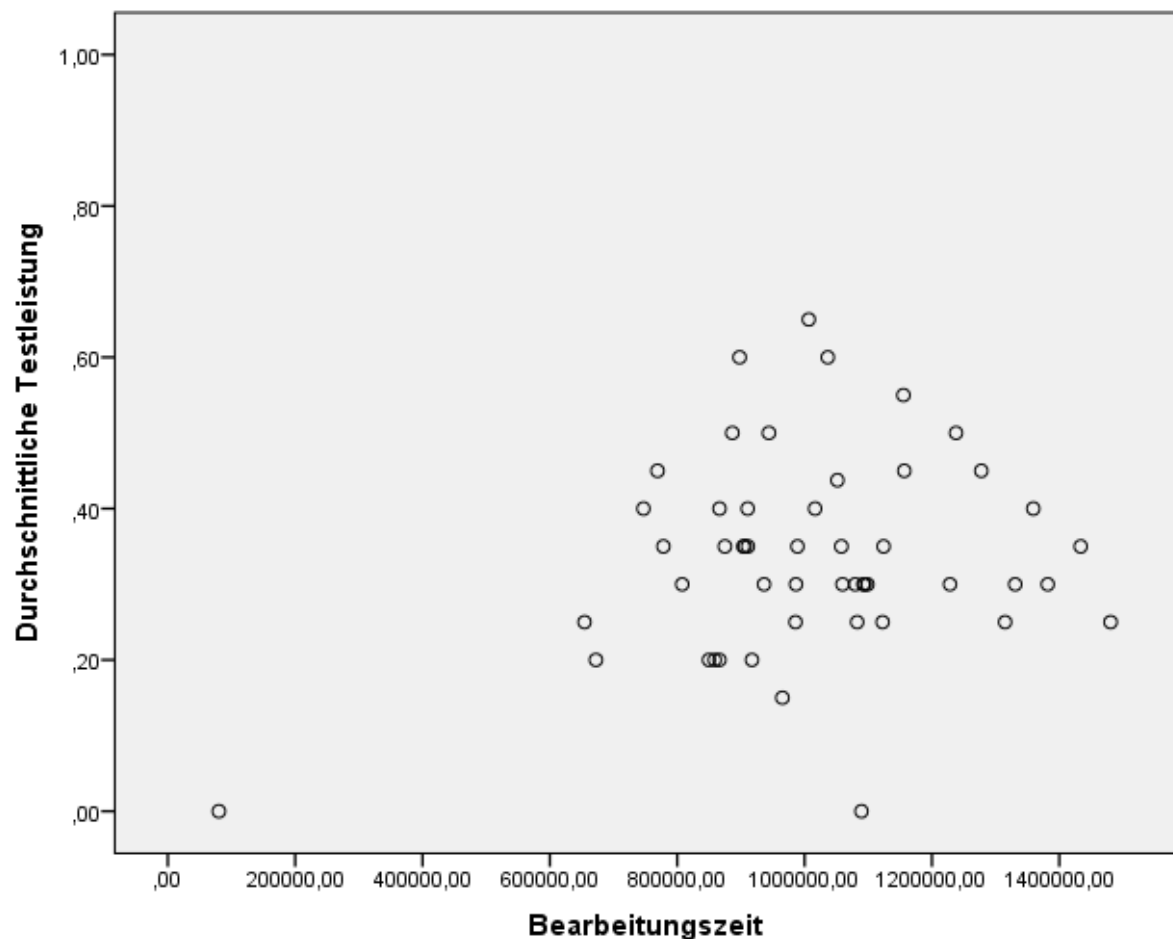


Abbildung 26: Streudiagramm zum Zusammenhang zwischen der kumulierten Bearbeitungszeit (in Millisekunden) und den durchschnittlichen Testleistungen der Probanden in der Skala zum prozeduralen politikdidaktischen Wissen (FpW)

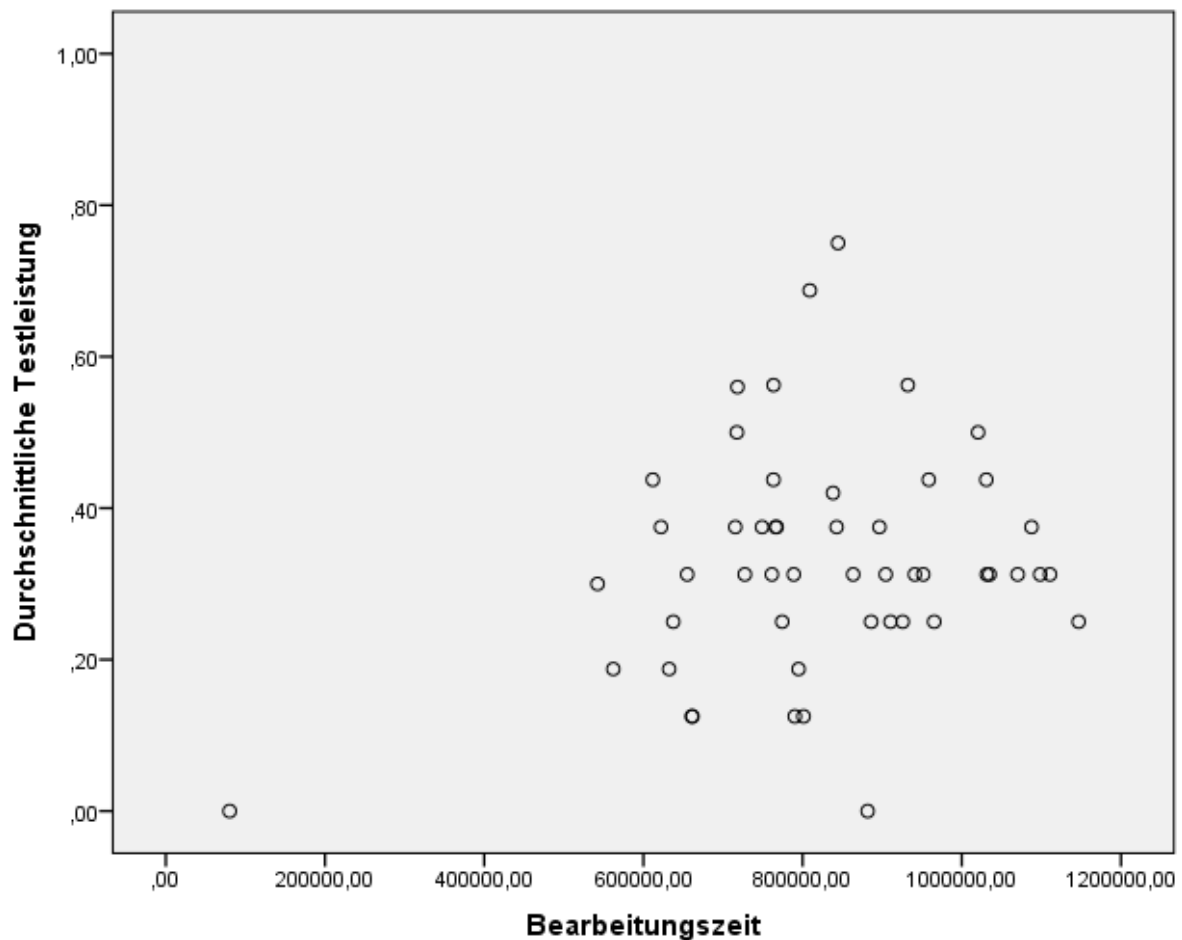


Abbildung 27: Streudiagramm zum Zusammenhang zwischen der kumulierten Bearbeitungszeit (in Millisekunden) und den durchschnittlichen Testleistungen der Probanden in der revidierten Skala zum prozeduralen politikdidaktischen Wissen (FpW_rev)

Die deutlich niedrige Bearbeitungszeit einer Probandin (Prob_ID 24) resultiert aus technischen Problemen im Computerprogramm, aufgrund derer nur zwei Items in der Skala zum prozeduralen politikdidaktischen Wissen bearbeitet wurden bzw. ausgewertet werden konnten. Da die Höhe einer Produkt-Moment-Korrelation deutlich von solchen Ausreißerwerten beeinflusst werden kann (Sedlmeier & Renkewitz, S. 225f.) und auch der Rangkorrelationskoeffizient (ρ) von Spearman anfällig gegenüber Ausreißern ist (Bortz & Lienert, 2008, S. 279), wird der Rangkorrelationskoeffizient von Kendall (τ) für die Berechnung herangezogen. Es zeigt sich eine sehr schwache, nicht signifikante Korrelation mit $\tau = .06$ ($p = .55$) zwischen der Bearbeitungszeit und der Testleistung in der Skala zum prozeduralen politikdidaktischen Wissen mit allen Items (FpW). Auch in der revidierten Skala zum prozeduralen

politikdidaktischen Wissen (FpW_rev) korreliert die Bearbeitungszeit der Probanden mit deren Testleistungen lediglich sehr schwach und nicht signifikant mit $\tau = .08$ ($p = .45$).

Zusammenfassend lässt sich sagen, dass die Testleistung der Probanden in den Skalen zum Fachwissen sowie zum deklarativen politikdidaktischen Wissen (alle Items) leicht überdurchschnittlich ist und die Leistungen in den jeweiligen revidierten Skalen im mittleren Bereich (FW_rev) sowie in einem leicht überdurchschnittlichen Bereich (FdW_rev) liegen. Die relativ niedrigen durchschnittlichen Testleistungen in den (revidierten) Skalen zum prozeduralen politikdidaktischen Wissen deuten darauf hin, dass die Probanden offenbar größere Schwierigkeiten bei der Lösung der Fragen hatten. Die hohe Bearbeitungszeit korreliert lediglich schwach mit den Testleistungen in der (revidierten) Skala zum prozeduralen politikdidaktischen Wissen. Der Befund, dass eine hohe Bearbeitungszeit keine Vorhersage einer hohen Testleistung erlaubt, ist konsistent zu den Erkenntnissen von Lindmeier (ebd., S. 173)⁵⁸ sowie zur Annahme, dass es den Studierenden unter Zeitdruck schwerfällt, adäquate Rückmeldungen zu geben.

7.2.6 Korrelationen

In der vorliegenden Arbeit interessieren auch die Zusammenhänge zwischen den Leistungen in den Skalen zu den einzelnen Strukturdimensionen, wobei Korrelationen nicht in einem kausalen Sinn interpretiert werden können (Bortz & Lienert, 2008, S. 258). Für die Berechnung der Zusammenhangsmaße wird im Folgenden der Rangkorrelationskoeffizient von Spearman (ρ) herangezogen (Bühner & Ziegler, 2009, S. 616ff.) (Tabelle 6).

⁵⁸ Wobei sich der Befund hierbei auf die Korrelation zwischen den Bearbeitungszeiten und Testleistungen von Lehrkräften bezieht.

Tabelle 6: Korrelationen zwischen den Skalen (Spearman's Rho)

Skala	FW	FpW
FW	-	$\rho = .34^* (p = .02)$
FdW	$\rho = .51^{**} (p < .001)$	$\rho = .51^{**} (p < .001)$
Skala	FW_rev	FpW_rev
FW_rev	-	$\rho = .38^{**} (p = .007)$
FdW_rev	$\rho = .49^{**} (p < .001)$	$\rho = .47^{**} (p < .001)$

** Die Korrelation ist auf dem .01 Niveau (zweiseitig) signifikant.

Es zeigt sich hierbei ein deutlicher, hoch signifikanter Zusammenhang zwischen dem Fachwissen (FW) und dem deklarativen politikdidaktischen Wissen (FdW) mit $\rho = .51$ ($p < .001$). Der Zusammenhang zwischen dem Fachwissen (FW) und dem prozeduralen politikdidaktischen Wissen (FpW) ist mäßig und signifikant mit $\rho = .34$ ($p = .02$) und der Zusammenhang zwischen dem deklarativen politikdidaktischen Wissen (FdW) und dem prozeduralen politikdidaktischen Wissen (FdW) fällt deutlich und hoch signifikant aus ($\rho = .51$; $p < .001$). In den revidierten Skalen zeigt sich ein moderater, hoch signifikanter Zusammenhang zwischen dem Fachwissen (FW_rev) und dem deklarativen politikdidaktischen Wissen (FdW_rev) mit $\rho = .49$ und $p < .001$. Zwischen dem Fachwissen (FW_rev) und dem prozeduralen politikdidaktischen Wissen (FpW_rev) zeigt sich eine mäßige, hoch signifikante Korrelation ($\rho = .38$; $p = .007$) und zwischen dem deklarativen politikdidaktischen Wissen (FdW_rev) und dem prozeduralen politikdidaktischen Wissen (FpW_rev) fällt der Zusammenhang moderat, hoch signifikant aus mit $\rho = .47$ und $p < .001$. Für den Zusammenhang der operationalisierten Strukturkomponenten lässt sich festhalten, dass das Fachwissen mit dem deklarativen politikdidaktischen Wissen etwas stärker korreliert als mit dem prozeduralen politikdidaktischen Wissen, das wiederum etwas deutlichere Zusammenhänge zum deklarativen politikdidaktischen Wissen zeigt (vgl. Kapitel 7.3 für eine Diskussion dieser Befunde).

Es kann schließlich vermutet werden, dass die Testleistungen in den Skalen mit Merkmalen der Probanden zusammenhängen. In Tabelle 7 und Tabelle 8 sind die durchschnittlichen Testleistungen in den (revidierten) Skalen in (Extrem-)Gruppen⁵⁹ gegenübergestellt.

Tabelle 7: Übersicht über die durchschnittlichen Testleistungen (Personenfähigkeit) in den Skalen sowie im Gesamttest gruppiert nach Personenmerkmalen

Testleistung	N	FW		FdW		FpW		Alle	
		M	SD	M	SD	M	SD	M	SD
männlich	21	.71	.09	.67	.11	.38	.12	.56	.19
weiblich	29	.55	.12	.60	.14	.31	.13	.49	.18
Abschluss Bachelor of Arts	26	.63	.14	.64	.14	.33	.13	.53	.20
Abschluss Erste Staatsprüfung	24	.60	.13	.63	.13	.34	.14	.52	.18
Schulform HRGe	12	.53	.15	.52	.13	.29	.07	.45	.16
Schulform GyGe	38	.64	.12	.67	.11	.35	.14	.55	.19
Unterrichtserfahrung nur im Praktikum	10	.59	.10	.63	.12	.28	.16	.50	.20
Unterrichtserfahrung (auch neben-)beruflich	39	.61	.14	.63	.13	.35	.12	.53	.18
Unterrichtserfahrung hoch (≥ 24 Std.)	14	.63	.12	.64	.13	.36	.13	.55	.18
Unterrichtserfahrung niedrig (≤ 6 Std.)	17	.61	.14	.62	.13	.35	.15	.53	.19

⁵⁹ Die Unterteilung in die Merkmalsausprägungen *Unterrichtserfahrung hoch* und *Unterrichtserfahrung niedrig* wird durch Rangbildung in drei Perzentile vorgenommen; das äußere Rangpaar bilden die beiden Extrema. Wegen der in der explorativen Datenanalyse erkennbaren Ausreißer im Merkmal *Unterrichtserfahrung hoch* (Abbildung A. 1 im Anhang) wird zur Prüfung, ob sich die so gebildeten Gruppen in ihren durchschnittlichen Testleistungen voneinander unterscheiden, der Mediantest herangezogen (Bortz & Lienert, 2008, S. 137ff.); der Unterschied zwischen den Gruppen ist hoch signifikant ($p < .001$).

Tabelle 8: Übersicht über die durchschnittlichen Testleistungen (Personenfähigkeit) in den revidierten Skalen sowie im revidierten Gesamttest gruppiert nach Personenmerkmalen

Testleistung	<i>N</i>	FW_rev		FdW_rev		FpW_rev		Alle_rev	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
männlich	21	.62	.15	.73	.18	.39	.14	.68	.20
weiblich	29	.42	.25	.59	.20	.29	.14	.43	.23
Abschluss Bachelor of Arts	26	.55	.25	.64	.20	.32	.13	.50	.24
Abschluss Erste Staatsprüfung	24	.54	.27	.65	.21	.34	.17	.51	.25
Schulform HRGe	12	.39	.26	.49	.21	.29	.09	.39	.21
Schulform GyGe	38	.59	.24	.69	.18	.34	.16	.54	.24
Unterrichtserfahrung nur im Praktikum	10	.50	.20	.59	.18	.28	.16	.46	.22
Unterrichtserfahrung (auch neben-)beruflich	39	.55	.27	.65	.20	.34	.15	.51	.25
Unterrichtserfahrung hoch (≥ 24 Std.)	14	.60	.25	.68	.21	.35	.15	.54	.25
Unterrichtserfahrung niedrig (≤ 6 Std.)	17	.52	.22	.63	.20	.35	.17	.50	.23

Die Vergleiche der (Extrem-)Gruppen werden mithilfe des U-Tests vom Mann-Whitney vorgenommen (Bortz & Lienert, 2008, S. 140). Sofern „der U-Test signifikant [wird], ist davon auszugehen, dass sich die Mediane der zugrunde liegenden Populationen unterscheiden (H_1)“ (ebd.). Tabelle 9 und Tabelle 10 zeigen die Ergebnisse des U-Tests (korrigiert für Rangbindungen) für die in Tabelle 7 und Tabelle 8 aufgeführten Merkmalspaare; die Effektstärke Phi in der Erläuterung ausgewählter Befunde wird anhand einer Formel approximiert.⁶⁰

⁶⁰ $\phi = \frac{z}{\sqrt{N}}$, wobei der ermittelte z -Wert durch die Stichprobengröße (N) geteilt wird (Bühner & Ziegler, 2009, S. 269). Die approximative Einschätzung der Effektstärke orientiert sich an den Konventionen $w = .10$ (kleiner Effekt), $w = .30$ (moderater Effekt) und $w = .50$ (starker Effekt); es wird dabei davon ausgegangen, „dass Φ und w eigentlich gleich sind“ (ebd., S. 270).

Tabelle 9: Ergebnisse des U-Tests auf Unterschiede in der zentralen Tendenz zwischen (Extrem-)Gruppen in den Skalen und im Gesamttest

	FW		FdW		FpW		Alle	
	U	<i>p</i>	U	<i>p</i>	U	<i>p</i>	U	<i>p</i>
männlich/weiblich	95.0	.00	211.0	.06	219.0	.09	1880.5	.00
Abschluss Bachelor of Arts/Erste Staatsprüfung	270.5	.42	282.5	.56	284.5	.59	2690.5	.66
Schulform HRGe/ GyGe	132.5	.03	89.5	.00	145.0	.06	1334.5	.00
Unterrichtserfahrung nur im Praktikum/(auch neben-)beruflich	171.0	.55	194.0	.98	167.5	.49	1645.0	.57
Unterrichtserfahrung hoch (≥ 24 Std.)/niedrig (≤ 6 Std.)	100.5	.46	106.5	.62	113.5	.83	986.0	.51

Tabelle 10: Ergebnisse des U-Tests auf Unterschiede in der zentralen Tendenz zwischen (Extrem-)Gruppen in den revidierten Skalen und im revidierten Gesamttest

	FW_rev		FdW_rev		FpW_rev		Alle_rev	
	U	<i>p</i>	U	<i>p</i>	U	<i>p</i>	U	<i>p</i>
männlich/weiblich	102.0	.00	189.5	.02	194.0	.03	1613.0	.00
Abschluss Bachelor of Arts/Erste Staatsprüfung	312.0	1.00	310.5	.98	293.5	.72	2780.5	.92
Schulform HRGe/ GyGe	125.0	.02	109.5	.01	175.0	.22	1299.0	.00
Unterrichtserfahrung nur im Praktikum/(auch neben-)beruflich	167.0	.49	158.0	.35	179.0	.69	1575.0	.39
Unterrichtserfahrung hoch (≥ 24 Std.)/niedrig (≤ 6 Std.)	88.5	.22	100.0	.45	114.0	.84	963.5	.41

Es zeigen sich Unterschiede in den Leistungen der *weiblichen und männlichen Probanden*. Die durchschnittlichen Testleistungen in allen (revidierten) Skalen sowie im jeweiligen Gesamttest sind bei den männlichen Probanden höher als bei den weiblichen. Diese Unterschiede sind in der unrevidierten Skala zum Fachwissen (FW), in der entsprechenden Gesamtskala sowie in allen drei revidierten Subskalen sowie der entsprechenden Gesamtskala signifikant; der approximierte Effekt ist hierbei für die Skalen FW ($\Phi = .58$), FW_rev ($\Phi = .57$) sowie für den revidierten Gesamttest Alle_rev ($\Phi = .61$) jeweils stark, für die Skalen FdW_rev ($\Phi = .32$), FpW_rev ($\Phi = .31$) sowie für den unrevidierten Gesamttest Alle ($\Phi = .46$) moderat bis stark und in den Skalen FdW ($\Phi = .26$) und FpW ($\Phi = .24$) klein bis moderat. In den Skalen sind die Differenzen zwischen den weiblichen und männlichen Probanden im prozeduralen politikdidaktischen Wissen (FpW) am geringsten ausgeprägt.

Zwischen den Probanden getrennt nach angestrebtem Studienabschluss (Bachelor of Arts bzw. Erste Staatsprüfung) zeigen sich keine signifikanten Unterschiede in den durchschnittlichen Testleistungen bei durchgängig kleinen approximierten Effektstärken. Dies kann als weitere Bestätigung für die Klassierung der Studierendengruppen zu einer Gesamtgruppe angesehen werden (vgl. Kapitel 7.1).

Statistisch signifikante Unterschiede zeigen sich wiederum in den durchschnittlichen Testleistungen der Probanden, die den Studiengang für Haupt-, Real- sowie die entsprechenden Jahrgangsstufen der Gesamtschulen studieren, und den Probanden, die den Studiengang für Gymnasien und Gesamtschulen studieren. Die Testleistungen der GyGe-Studierenden liegen in allen revidierten und unrevidierten Skalen auf höherem Niveau als die Testleistungen der HRGe-Studierenden, wobei die Unterschiede in allen Skalen mit Ausnahme der Skala zum prozeduralen politikdidaktischen Wissen sowie den entsprechenden Gesamtskalen signifikant sind. Die approximierten Effekte sind für die Skalen FW ($\Phi = .30$), FW_rev ($\Phi = .33$), FdW ($\Phi = .45$) sowie FdW_rev ($\Phi = .38$) moderat bis stark und für die Skalen FpW ($\Phi = .27$), FpW_rev ($\Phi = .17$) sowie für die (unrevidierten) Gesamttests Alle ($\Phi = .26$) und Alle_rev ($\Phi = .27$) klein bis moderat.

Testpersonen, die angegeben haben, bereits eigene Unterrichtserfahrungen neben den studienbegleitenden Praktika gesammelt zu haben, erreichten in allen Skalen – bis auf die Skala FdW – und in den beiden Gesamttests etwas höhere Testleistungen als Probanden, die zum Zeitpunkt der Erhebung Unterrichtserfahrungen ausschließlich in studienbegleitenden Prak-

tika gesammelt hatten; die Unterschiede sind jedoch nicht signifikant bei durchgängig kleinen Effektstärken.

Die Testleistungen von Probanden, die – relativ zu den Angaben aller Teilnehmenden – umfangreichere Unterrichtserfahrungen gesammelt haben, sind in allen Skalen mit Ausnahme der Skala FpW_rev sowie in den jeweiligen Gesamtscores etwas höher als die Leistungen derjenigen mit kaum oder wenigen Unterrichtserfahrungen; wiederum sind die Unterschiede nicht signifikant bei durchgängig kleinen Effektstärken.

7.3 Zusammenfassung und kritische Reflexion

Die nachfolgende Zusammenschau fokussiert die zentralen empirischen Befunde unter Bezug auf die üblichen Gütekriterien, wie sie grundständig in Kapitel 2.3.3.1 ausgeführt wurden, und reflektiert diese, wie auch die Ergebnisse zu den erreichten Testleistungen der Probanden, kritisch.

Objektivität

Um die *Durchführungsobjektivität* für die vorliegende Studie sicherzustellen, wurde vorab eine standardisierte Vorgehensweise in einem Verlaufsskript festgelegt. In diesem sind die genauen Testinstruktionen wörtlich festgelegt, um die Durchführung der Studie von möglicherweise inkonsistentem Verhalten des Versuchsleiters unabhängig zu halten und um Rückfragen seitens der Probanden zu minimieren (Sedlmeier & Renkewitz, 2008, S. 71). Der Ablauf des Tests wurde schriftlich in einem Verlaufsprotokoll festgehalten, um ggf. auftretende Störungen oder Unregelmäßigkeiten dokumentieren zu können. Die Bedingungen der Testdurchführungen wurden dabei – wo forschungspraktisch möglich – gleich gehalten. Eine Standardisierung des Tests (Schmidt-Atzert & Amelang, 2012, S. 132) wurde durch die Implementation von Videovignetten in einem Computerprogramm angestrebt. Die Videovignetten selbst, aber auch die Aufgabenformate blieben dabei für alle Probanden gleich, was für einen hohen Standardisierungsgrad spricht (Lindmeier, 2013, S. 53).

Da ein Großteil der in dieser Studie verwendeten Items geschlossen formuliert ist und die Antworten dementsprechend in der Auswertung eindeutig bewertbar sind, kann für diese

Testbestandteile von einer hohen *Auswertungsobjektivität* ausgegangen werden (Bortz & Döring, 2006, S. 195). Gleiches gilt theoretisch auch für die im Test enthaltenen halb offenen Fragen, da diese so formuliert wurden, dass jeweils nur eine Lösung richtig ist (ebd., S. 213). Um die Auswertungsobjektivität für diese Antworttypen dabei möglichst vollständig sicherzustellen, wurde für die halb offenen Items ein Ratingmanual verfasst, das von zwei Experten unabhängig angewendet wurde, um die Antworten zu bewerten (consensus expert scoring). Offene Antworten können die Auswertungsobjektivität beeinträchtigen (Sedlmeier & Renkewitz, 2008, S. 71), weshalb auch für die Bewertung der offenen Antworten in der Skala zum prozeduralen politikdidaktischen Wissen ein Ratingmanual erstellt wurde, das die zu beurteilenden Kategorien genau definiert, Ankerbeispiele für die Bewertung enthält und Differenzierungen für die Begutachtung der beiden Antworttypen beschreibt. Die Antworten der Probanden wurden von zwei geschulten Ratern unabhängig voneinander beurteilt. Die Güte der Beurteilungsübereinstimmung (Urteilskonkordanz) lässt sich in der vorliegenden Studie aufgrund durchweg guter Konkordanzwerte hoch einstufen. Trotz des offenen Antwortformats in der Skala zum prozeduralen politikdidaktischen Wissen kann aufgrund der Ratinganleitungen von einem hohen Standardisierungsgrad ausgegangen werden (Lindmeier, 2013, S. 53).

Um eine möglichst hohe *Interpretationsobjektivität* für die offenen Items zum prozeduralen politikdidaktischen Wissen zu sichern, fand die Bewertung der Antworten der Probanden auf einer mehrstufigen Skala statt. Die hohe Urteilskonkordanz kann dabei auch als Zeichen für eine hohe Interpretationsobjektivität in dieser Skala gewertet werden. Aufgrund zum Teil weiterhin ausstehender domänenspezifischer Forschungen sind dabei kriteriale Referenzpunkte für eine *absolute* Interpretation der Ergebnisse der Studie schwierig auszumachen. Der Umstand, dass in der vorliegenden Studie eine Gelegenheitsstichprobe untersucht wurde, macht es zudem notwendig, auch die *relativen* Testleistungen der Probanden vorsichtig zu interpretieren, da sich die Normalverteilung der Testleistungen (Personenfähigkeit) auf den jeweiligen Wertebereich bezieht, innerhalb dessen alle Leistungen verteilt sind. Verdeutlicht werden kann dies an der Skala zum prozeduralen politikdidaktischen Wissen: Die Testwerte der Probanden in dieser Fragebatterie können nur für diese Untersuchung als niedrig eingestuft werden, da entsprechende Bezugspunkte ‚vom anderen Ende der Skala‘ – bspw. von Personen mit sehr hoher Lehrexpertise – derzeit noch nicht vorliegen. Dem kann entgegengehalten werden, dass sich der explorative Zweck der vorliegenden Studie zunächst auf

die Erkundung möglicher Ansätze zur Weiterentwicklung bestehender Instrumente zur Erhebung politikdidaktischen Wissens erstreckt und insofern Rückschlüsse auf eine Grundgesamtheit zunächst nicht von Interesse sind. Zudem zeigen sich die vorliegenden Interpretationen – auch zum prozeduralen politikdidaktischen Wissen – grundsätzlich kongruent zu den theoretischen Annahmen sowie zu den empirischen Befunden aus anderen Domänen, sodass von einer für die hier notwendigen Zwecke hohen Interpretationsobjektivität ausgegangen werden kann.

Reliabilität

Die Reliabilität wurde in der vorliegenden Studie durch den Alphakoeffizienten von Cronbach als Maß der *internen Konsistenz* überprüft, deren Berechnung als stabile Schätzung der Reliabilität angesehen wird (Bortz & Döring, 2006, S. 198). Die hierzu zunächst durchgeführte Itemanalyse zeigte in den unrevidierten Skalen mit allen Items, in den Skalen zum Fachwissen sowie zum deklarativen politikdidaktischen Wissen einige Items mit negativer *Itemtrennschärfe*, die sich wiederum auf die Alpha-Werte auswirken können. Hiermit zusammenhängend konnten hinsichtlich der *Itemschwierigkeit* Deckeneffekte in den Skalen zum Fachwissen und zum deklarativen politikdidaktischen Wissen sowie Bodeneffekte in der Skala zum prozeduralen politikdidaktischen Wissen identifiziert werden. Die Zusammenhänge zwischen Schwierigkeit und Trennschärfe der Items konnten am Beispiel der Skala zum Fachwissen aufgezeigt werden. Es wurde begründet, warum jedoch eine vorschnelle Eliminierung von Items, die die Analyse als nicht passend auswies, zum Zwecke der Erhöhung der Reliabilität unter Umständen negative Auswirkungen auf die Validität des Tests haben könnte. Es wurden daher nur diejenigen Items von der weiteren Datenanalyse probe-weise ausgeschlossen, die sich auch inhaltlich als weniger passend charakterisieren lassen (Bühner, 2006, S. 106f.). Zudem wurden Items behalten, die auch in den Randbereichen differenzieren (ebd., S. 99; Schmidt-Atzert & Amelang, 2012, S. 114f.).

Unter Ausschluss unpassender Items zeigt die Skala zum Fachwissen schließlich einen guten Alpha-Wert ($\alpha = .85$); die Reliabilität der revidierten Skalen zum deklarativen ($\alpha = .63$) sowie zum prozeduralen politikdidaktischen Wissen ($\alpha = .61$) können vor dem Hintergrund der relativ niedrigen Itemanzahl sowie unter Berücksichtigung der explorativen Zwecke als noch akzeptabel betrachtet werden. Der Vergleich der Reliabilität der revidierten Skalen mit der

internen Konsistenz der Gesamtskala ($\alpha = .70$) kann als Anhaltspunkt dafür angesehen werden, dass Letztere kein generelles Konstrukt erfasst (Lindmeier, 2011, S. 168). Das verdeutlicht auch die gepaarte Berechnung der internen Konsistenz der revidierten Skalen, die zeigt, dass die Reliabilität niedriger ist, wenn die revidierte Skala zum prozeduralen politikdidaktischen Wissen hinzugenommen wird (FW_rev UND FpW_rev: $\alpha = .66$; FdW_rev UND FpW_rev: $\alpha = .20$). Wird berücksichtigt, dass die Testleistungen der Probanden (Personenfähigkeit) im Fachwissen und im prozeduralen politikdidaktischen Wissen lediglich mäßig, hoch signifikant miteinander korrelieren ($\rho = .38$; $p = .007$) und der Zusammenhang der Testleistungen in den revidierten Skalen zum deklarativen politikdidaktischen Wissen sowie zum prozeduralen politikdidaktischen Wissen moderat, hoch signifikant ausfällt ($\rho = .47$ und $p < .001$), kann davon ausgegangen werden, dass die die Skala zum prozeduralen politikdidaktischen Wissen entlang der zugrunde liegenden Modellierung des politikdidaktischen Wissens in der vorliegenden Studie eine separate Strukturdimension ist. Da die Reliabilität der probeweise zusammengefassten revidierten Skalen zum Fachwissen und zum deklarativen politikdidaktischen Wissen im Vergleich mit der internen Konsistenz der jeweiligen Subskala nicht sinkt ($\alpha = .85$) und die Testleistungen in den beiden Subskalen moderat, hochsignifikant ($\rho = .49$ und $p < .001$) korrelieren, ist konsistent zur Modellierung und Operationalisierung des politikdidaktischen Wissens in der vorliegenden Studie, da das lernprozessbezogene deklarative politikdidaktische Wissen durch profundes Fachwissen gekennzeichnet ist. Kritisch zu diskutieren ist jedoch, ob dann die so konstruierten Items tatsächlich einen Aspekt des fachdidaktischen Wissens der Probanden erfassen (vgl. auch Abbildung 5 und Abbildung 21) oder ob für die korrekte Beantwortung der Fragen entsprechendes Fachwissen ausreicht. Hierauf sollte bei zukünftigen Itemkonstruktionen zum lernprozessbezogenen deklarativen politikdidaktischen Wissen genauer eingegangen werden. Dies würde auch im Ansatz erklären, warum die Testleistungen in den revidierten Skalen zum Fachwissen und zum prozeduralen politikdidaktischen Wissen – unter der theoretischen Annahme, dass das Vorhandensein lernprozessbezogenen Wissens auch in dieser Strukturdimension bedeutsam wird (vgl. Kapitel 6.3.1.2) – lediglich mäßig, hoch signifikant ($\rho = .38$; $p = .007$) miteinander korrelieren und die Genauigkeit der Messung der paarweise gebildeten Skala aus den revidierten Skalen zum deklarativen sowie prozeduralen politikdidaktischen Wissen wesentlich niedriger ist als die entsprechenden einzelnen revidierten Subskalen. Offenbar stellen die Anforderungssituationen in den Test-Items zum prozeduralen politikdidaktischen Wissen die Probanden vor Herausforderungen, für deren

Bewältigung des Vorhandenseins deklarativen Wissens nicht ausreicht. Aufgrund des explorativen Charakters der Studie ist die Vorläufigkeit dieser Annahmen jedoch nochmals zu betonen.

Validität

In der folgenden Betrachtung stehen die *Inhaltsvalidität* und die *Konstruktvalidität* im Fokus.⁶¹ Das heuristische Strukturmodell des politikdidaktischen Wissens diente in dieser Arbeit auch der Entwicklung von Test-Items. Hierzu wurde ein höherer Auflösungsgrad der Modellierung vorgeschlagen, der durch die Strukturdimension der Inhaltsbereiche die zunächst sehr große Menge an Items, die für eine Operationalisierung der modellierten Konstrukte infrage kämen, deutlicher eingrenzbar machte. Wie bereits angesprochen, wurde zudem bei der Itemanalyse beachtet, dass unpassende Items von den weiteren Analyseschritten nur dann probeweise ausgeschlossen wurden, wenn sie inhaltlich ebenfalls weniger geeignet waren (Bühner, 2006, S. 106f.). Hierdurch sollte die *Inhaltsvalidität* als gegeben angesehen werden können.

Für die *diskriminante Validität* als Teilaspekt der *Konstruktvalidität* betonen Krauss et al. (2011) zur Testkonstruktion von COACTIV im Hinblick auf die Korrelationen zwischen den Strukturkomponenten:

„Die Tatsache, dass die Korrelationen sich aber dennoch deutlich von 1 unterscheiden, ist ein erster Hinweis auf die diskriminante Validität des Tests, das heißt, trotz des starken Zusammenhangs messen die verschiedenen Tests (bzw. die Subfacetten des fachdidaktischen Wissens) auch tatsächlich etwas Verschiedenes.“ (S. 150)

Ähnliches deutet sich für die Korrelationen zwischen den Strukturkomponenten in der vorliegenden Studie an (vgl. auch Tabelle 6). Ausgeprägte Zusammenhänge zwischen den Skalen zeigen sich durchgängig nicht, sodass davon ausgegangen werden kann, dass die Skalen etwas Unterschiedliches messen. Gleichzeitig stützen die reliabilitätsstatistischen

⁶¹ Wie bereits in Kapitel 2.3.3 angesprochen, ist die Identifikation entsprechender Außenkriterien für die Bestimmung der Güte der Kriteriumsvalidität schwierig. Für die Stichprobe wäre es z. B. denkbar, als Außenkriterium für die Überprüfung der Validität der Skalen zum Fachwissen sowie zum deklarativen politikdidaktischen Wissen entsprechende Prüfungsleistungen im Studium heranzuziehen; diese Möglichkeit besteht jedoch schon aus Gründen des Datenschutzes nicht. Entsprechende Außenkriterien zur Überprüfung der Skala zum prozeduralen politikdidaktischen Wissen zu finden, dürfte schließlich noch schwieriger sein.

Analysen die theoretische Annahme, dass die einzelnen Strukturkomponenten des politikdidaktischen Wissens dennoch miteinander verbunden sind (Weschenfelder, 2014a, S. 190). Vor allem aber das prozedurale politikdidaktische Wissen lässt sich deutlicher von den anderen Wissenskomponenten unterscheiden. Es ist hierbei anzumerken, dass bei der Messung einer solchen relativ komplexen Strukturkomponente das gemessene Konstrukt mit weiteren Merkmalen konfundiert sein kann, was sich mindernd auf die diskriminante Validität auswirken kann. So wäre es möglich, dass auch Facetten pädagogisch-psychologischen Wissens miterfasst werden.

Testschwierigkeit und Testleistung

Die Analyse der Test-Items und der durchschnittlichen Leistungen der Probanden (Personenfähigkeit) in den Skalen zeigt, dass sich diese in ihrem Schwierigkeitsgrad unterscheiden. In den revidierten Skalen liegen die Testwerte im Fachwissen auf einem durchschnittlichen Niveau ($M = .54$; $SD = .26$) und im deklarativen politikdidaktischen Wissen auf einem leicht überdurchschnittlichen Niveau ($M = .65$; $SD = .20$). Die Testleistungen in der revidierten Skala zum prozeduralen politikdidaktischen Wissen sind deutlich unterdurchschnittlich ($M = .33$; $SD = .15$); die Streuung der Testwerte ist in dieser Skala geringer als in den beiden anderen Skalen. Die durchschnittlichen Testleistungen in der revidierten Fragebatterie zum prozeduralen politikdidaktischen Wissen korrelieren dabei sehr schwach, nicht signifikant mit den kumulierten Bearbeitungszeiten der Probanden ($\tau = .08$; $p = .45$). Die insgesamt niedrigeren durchschnittlichen Testleistungen der Probanden in der Skala zum prozeduralen politikdidaktischen Wissen verbunden mit dem Befund, dass auch eine höhere Bearbeitungszeit nur sehr schwach mit diesen Leistungen korreliert, sind konsistent zu den theoretischen Annahmen, dass sich prozedurales fachdidaktisches Wissen aus dem reflektierten Erfahrungswissen von Lehrpersonen speist. Von einem solchen Wissen kann bei den Studierenden nur bedingt ausgegangen werden.

Die vergleichende Analyse der Leistungen der Probanden getrennt nach verschiedenen Personenmerkmalen zeigt, dass die männlichen Probanden durchgängig höhere Testleistungen erbracht haben als die weiblichen. Die statistisch signifikanten Unterschiede zwischen den HRGe-Studierenden und den GyGe-Studierenden fallen insoweit erwartungsgemäß aus, als sich ähnliche Unterschiede auch in der PKP-Studie zeigten (Weißeno et al., 2013, S. 191).

Denkbar ist jedoch auch, dass dieses Merkmal mit dem Personenmerkmal *Geschlecht* konfundiert ist, da zwei Drittel der getesteten HRGe-Studierenden weiblich sind, während dies nur etwa auf die Hälfte der getesteten GyGe-Studierenden zutrifft.

Schließlich wurde erwartet, dass Probanden mit umfangreicherem unterrichtlichem Erfahrungswissen in der Skala zum prozeduralen politikdidaktischen Wissen bessere Testleistungen zeigen würden, als diejenigen, die bis zum Zeitpunkt des Tests kaum oder wenige eigene Unterrichtserfahrungen gesammelt hatten. Die empirischen Befunde stützen diese Annahme nicht, wobei zu berücksichtigen ist, dass die Differenzen zwischen den Merkmalsausprägungen *umfangreiche versus weniger umfangreiche Unterrichtserfahrungen* sicher geringer ausgeprägt sind, als sie bspw. zwischen Studierenden mit wenig Erfahrungswissen versus Lehrkräften mit langjähriger Berufserfahrung wären.

Dass die Stichprobe in dieser Studie nicht alle Charakteristika einer Eichstichprobe aufweist (Bortz & Döring, 2006, S. 217), sollte bei der Betrachtung der Testschwierigkeit sowie der Testleistungen berücksichtigt werden, was Schmidt-Atzert und Amelang (2012) für das gewählte methodische Vorgehen nach der klassischen Testtheorie (*KTT*) so ausdrücken: „Die Itemschwierigkeiten sind ohnehin offensichtlich von der untersuchten Population oder Stichprobe abhängig: Je stärker das Merkmal bei den untersuchten Personen ausgeprägt ist, desto höher fallen die Itemschwierigkeiten aus“ (S. 59). Aufgrund der Ziele des Studiums – der Aufbau insbesondere deklarativen (theoretisch-formalen) Wissens – kann vermutet werden, dass es den Probanden leichter fällt, Fragen zu Fakten und Sachverhalten der Politikdidaktik zu beantworten, als anwendungsbezogene Aufgabenstellungen zu bearbeiten, und dass daher die Testleistungen in den Skalen zum deklarativen Wissen höher ausfallen. Hinzu kommt, dass zwar alle Studierenden im Lehramt Sozialwissenschaften die Möglichkeit hatten, an der Studie teilzunehmen, sofern sie sich bereits in einem bestimmten Studienabschnitt befanden, die Stichprobenziehung jedoch nicht zufällig erfolgte. Daher kann grundsätzlich nicht ausgeschlossen werden, dass sich die so gewonnenen Probanden der Stichprobe in testleistungsrelevanten Merkmalen von der Grundgesamtheit unterscheiden (ebd.). So ist es möglich, dass sich – in Bezug auf die hier getesteten Merkmale – eher ‚leistungsstärkere‘ Personen zur Teilnahme gemeldet haben. Überprüft werden könnte dies z. B. anhand eines Vergleichs der Testwerte mit den Studienleistungen der Versuchspersonen. Diese Möglichkeit entfällt jedoch aus datenschutzrechtlichen Gründen.

Die kritische Reflexion der empirischen Befunde zeigt Grenzen beim methodischen Vorgehen nach der KTT auf. Wie Hartig und Frey (2013) fordern, sollten „[z]ur Messung und psychometrischen Modellierung von Kompetenzen (vgl. z. B. Hartig 2008) ... vorzugsweise Methoden mit latenten Variablen (vgl. z. B. Skrondal und Rabe-Hesketh 2004) zum Einsatz kommen“ (S. 48), wozu „Modelle der Item-Response-Theorie (IRT), je nach Datenlage und Zielsetzung aber auch andere Modellklassen wie beispielsweise Strukturgleichungsmodelle [zählen]“ (ebd.; vgl. auch Hartig & Höhler, 2010, S. 190). Probabilistische Modelle der IRT haben gegenüber der KTT einige Vorteile – beispielhaft seien Möglichkeiten der empirischen Überprüfbarkeit der Modellpassung, die Stichprobenunabhängigkeit (Hartig, 2009, S. 298) oder auch erweiterte Strategien zur Konstruktion von Kompetenzniveaus (Hartig & Jude, 2007, S. 32ff.) genannt –, konnten wegen des explorativen Charakters der vorliegenden Arbeit und der damit verbundenen Anlage der Studie jedoch nicht zur Anwendung kommen (Bortz & Döring, 2006, S. 221; Lindmeier, 2011, S. 177; zur Notwendigkeit einer adäquaten Stichprobengröße vgl. auch Hartig & Frey, 2013, S. 50). In Folgestudien sollte ein methodisches Vorgehen auf Basis der IRT erfolgen.

8 Diskussion und Ausblick

In diesem Kapitel werden die zentralen Ergebnisse der Arbeit zusammengefasst und kritisch reflektiert (Kapitel 8.1). Auf dieser Grundlage werden abschließend Ansätze für die weiterführende Forschung sowie mögliche Implikationen für die Aus- und Weiterbildung von Politiklehrerinnen und Politiklehrern diskutiert (Kapitel 8.2).

8.1 Zusammenfassung und Schlussfolgerungen

Im ersten Teil der Arbeit lag der Fokus auf der Analyse der theoretischen Grundlagen sowie der bislang vorliegenden empirischen Erkenntnisse zu professionellen Kompetenzen von Lehrkräften aus einer politikdidaktischen, aber auch fächerübergreifenden Perspektive (*Erste Forschungsfrage*). Es konnte gezeigt werden, dass dem professionellen Wissen und Können von (angehenden) Lehrkräften im Fach Politik in der politikdidaktischen Forschung im Sinne eines kompetenztheoretischen Ansatzes bislang erst wenig Aufmerksamkeit gewidmet wurde, was insofern besonders überrascht, als der Expertise von Lehrerinnen und Lehrern eine herausragende Bedeutung für die Planung, Gestaltung und Evaluation unterrichtlicher Lehr-/Lern-Prozesse zukommt (Besser & Krauss, 2009; Bromme, 1992/2014; Schmelzing et al., 2010, S. 190; Tepner et al., 2012; Weißeno et al., 2013, S. 187). Zwar gab es vereinzelt Überlegungen zur professionellen Kompetenz von Lehrkräften (Petric, 2009, 2012; Reinhardt, 2009, 2012a), diese zeigten sich jedoch wenig anschlussfähig an kompetenztheoretische Grundlagen und enthielten auch keine explizite Modellierung der Struktur des Wissens und Könnens von Politiklehrerinnen und -lehrern. Erst das PKP-Forschungsprogramm (Oberle et al., 2012) hat ein kompetenztheoretisch fundiertes und an bestehende Forschungskontexte anderer Fächer anschlussfähiges Modell vorgelegt und empirisch validiert. Gleichzeitig, so geben auch die Autoren dieser Studie zu bedenken, müssen Forschungen zu den Kompetenzen von (angehenden) Politiklehrerinnen und Politiklehrern intensiviert werden, damit bestehende Forschungslücken systematisch geschlossen werden können (Weißeno et al., 2013, S. 200).

Die Zusammenschau dieser Forschungslücken im ersten Teil der vorliegenden Arbeit bestätigte die von Weißeno et al. (ebd.) skizzierten Weiterentwicklungsmöglichkeiten der Modellierung und Operationalisierung professioneller Kompetenzen und zeigte darüber

hinaus noch weitere Aspekte auf, durch die ein umfassendes Verständnis zum professionellen Wissen und Können von (angehenden) Politiklehrkräften erlangt werden kann. Dabei konnte als wesentliche Aufgabe zukünftiger Forschungen die Modellierung und Operationalisierung *handlungsnaher* Aspekte professioneller Kompetenzen von Politiklehrerinnen und -lehrern benannt werden. Wenn die professionelle Kompetenz von Lehrkräften (wie auch in den Studien COACTIV, MT21 und PKP) in schriftlichen Tests erfasst werden soll, wird dabei in der Regel davon ausgegangen, dass das so erfasste theoretisch-formale Wissen eine wesentliche Ausgangsbasis für erfolgreiches unterrichtliches Handeln darstellt (Fenstermacher, 1994). Die Annahme, dass in schriftlichen Leistungstests erfassbares deklaratives Wissen auch in eine adäquate Handlungsstrategie überführt werden kann, muss aus theoretischer Perspektive hinterfragt werden (Vogelsang & Reinhold, 2013). Riese und Reinhold (2010) formulieren dies so:

„Einigkeit besteht zunächst darin, dass deklaratives Wissen – etwa die Kenntnis von Konzepten und Fakten – allein in der Unterrichtspraxis nicht ausreicht und auch handlungsnahes, situationsorientiert organisiertes prozedurales Wissen vonnöten ist, welches sich beispielsweise in der Verfügbarkeit von Handlungsroutrinen äußert. Prozedurales (fachdidaktisches Wissen) kann jedoch nur bis zu einem gewissen Grad expliziert werden, da es insbesondere beim Handeln unter Zeitdruck immer einen Satz von Hintergrundannahmen gibt, die das Handeln leiten (Oberauer, 1993). Nicht zuletzt aufgrund unterschiedlicher theoretischer Perspektiven herrscht allerdings Unklarheit hinsichtlich der mentalen Repräsentation solcher Wissensbestände. Ein einfaches Transfermodell kann der komplexen Wechselwirkung zwischen Wissen und Handeln nicht gerecht werden, so haben Berufsanfänger Probleme bei der Umsetzung theoretischen Wissens in Handlung (Fischler, 2008).“ (S. 172)

Neben den angesprochenen theoretischen Grenzen sind hiermit gleichzeitig forschungspraktische Herausforderungen benannt, da Verfahren zur Diagnostik solch komplexer Konstrukte wie *politikdidaktischer Handlungskompetenz* in herkömmlicher schriftlicher Form nicht angemessen oder nur mit einem erheblichen Itemumfang erfassen können (Frey & Hartig, 2013, S. 55). Studien wie COACTIV und MT21, die bewusst keine handlungsnahen Aspekte professioneller Kompetenzen messen (Riese & Reinhold, 2010, S. 174), werden dementsprechend dahingehend problematisiert, dass sie unter anderem die Situationsspezifität und den Anforderungsreichtum des Unterrichts nicht adäquat berücksichtigen (Oser et al., 2010, S. 5f.). Auch für die Politikdidaktik wird der Bedarf einer deutlicheren Fokussierung auf den Zusammenhang von Wissen und Können formuliert (Weschenfelder, 2014a, S. 79). Im Rahmen der theoretischen Exploration der vorliegenden Arbeit wurden Weiterentwicklungsmöglichkeiten identifiziert, die bei der Schließung dieser Forschungslücken helfen können. So

scheint es angebracht, die Modellierung professioneller Handlungskompetenz um die Strukturdimension *Wissensarten* (deklarativ/prozedural) zu erweitern. Für die Modellierung des Professionswissens von Politiklehrkräften wurde eine Orientierung an dem von Tepner et al. (2012, S. 19) entwickelten mehrdimensionalen Modell begründet und es wurde hierauf aufbauend ein heuristisches Strukturmodell politikdidaktischen Wissens und Könnens vorgeschlagen, das dazu dienen soll, eine theoretisch begründete Ausgangsbasis für die Entwicklung von Test-Items bereitzustellen. Das Modell spezifiziert neben den Wissensarten auch die Strukturdimension *Inhaltsbereiche*. Dies wurde durch empirische Befunde aus anderen Fächerdomänen begründet, die darauf hindeuten, dass „hohes Wissen in einem Themengebiet nicht automatisch hohes Wissen in einem anderen Themengebiet impliziert“ (Tepner et al., 2012, S. 9). Da „sich das fachdidaktische Wissen auf die *inhaltsbezogenen* (nicht allgemein-methodischen) unterrichtlichen Aufgaben [bezieht]“ (Krauss et al., 2008b, S. 227) und die entsprechenden empirischen Befunde in anderen Fächern (für COACTIV vgl. Kunter & Baumert, 2011, S. 347) und in der Politikdidaktik darauf hindeuten, „dass es sich bei fachwissenschaftlichem und fachdidaktischem Wissen um zusammenhängende, aber trennbare Konstrukte handelt“ (Weschenfelder, 2014a, S. 222), ist begründbar, dass das politikdidaktische Wissen inhaltsbezogen erfasst werden kann. Eine Anschlussfähigkeit an Forschungen anderer Fächerdomänen (Tepner et al., 2012; van Dijk & Kattmann, 2010) sowie an die Modellierung professioneller Kompetenzen in der PKP-Studie ist weiterhin gegeben. Zudem wird hierdurch im Prinzip der Auflösungsgrad der Modellierung erhöht, was die ‚Füllung‘ der Facetten des politikdidaktischen Wissens erleichtern kann.

Zu diesem Zweck wurde in einer vergleichenden Gegenüberstellung analysiert, wie die Konstrukte in unterschiedlichen Studien operationalisiert werden; das Wissen über fachspezifische Instruktions- und Vermittlungsstrategien und das Wissen über (Fehl-)Vorstellungen der Lernenden konnten dabei als die beiden zentralen Facetten des fachdidaktischen Wissens benannt werden (Gramzow et al., 2013, S. 14; Schmelzing, 2010, S. 23; Tepner et al., 2012, S. 14; van Dijk & Kattmann, 2010, S. 8; Weschenfelder, 2014a, S. 109). Es wurde dabei problematisiert, dass weiterhin theoretische und empirische Defizite der Politikdidaktik (Weißeno, 2012, S. 124) erkennbar sind und eine sich auf theoretische Erkenntnisse sowie auf Forschungsbefunde stützende Begründung der Operationalisierung der Facetten des politikdidaktischen Wissen erschweren, da gefordert wird, dass „[d]ie Konstruktion eines Kompetenzmodells ... sich von Anfang an auf empirisch gestützte Theorien zu kognitiven

Prozessen beziehen [sollte]“ (Leuders & Sodian, 2013, S. 32). Dahingehend kann auch der in dieser Arbeit gewählte normativ-deduktive Zugang kritisch diskutiert werden (Blömeke, 2013b, S. 34). Ein solcher Zugang wäre idealerweise durch entsprechende qualitative und quantitative empirische Studien zu stützen. Jedoch ist zum Beispiel bei der Konzeptualisierung und Operationalisierung der Facetten des politikdidaktischen Wissens zu berücksichtigen, dass Forschungen zur Wirksamkeit von fachspezifischen Methoden und didaktischen Prinzipien weiterhin ausstehen (Manzel & Neumann, 2015, S. 77). Ähnliches lässt sich auch hinsichtlich der Auswahl und Begründung der Videovignetten als Stimulus für das video-cued testing in vorliegenden Studie feststellen: Zwar wurden die Videovignetten entlang der in dieser Arbeit interessierenden theoretisch begründeten Referenzpunkte zu Aspekten hoher Lehrerexpertise selektiert (vgl. Kapitel 2.3.3.3); wenn jedoch die so dargebotenen Stimuli nicht nur authentische Anforderungssituationen enthalten sollen – davon sollte hier aufgrund der Authentizität der Ausgangsmaterialien ausgegangen werden können –, sondern auch in „prototypische Unterrichtsszenarien einzubetten [sind]“ (Riese & Reinhold, 2014, S. 262), erfordert dies detaillierte Kenntnisse über *das Typische* von Politikunterrichtsstunden. Hierzu gibt es jedoch ebenso Forschungsbedarf (vgl. Kapitel 2.3.3.2) wie zu typischen (Fehl-)Vorstellungen der Lernenden im Fach Politik (vgl. Kapitel 2.2.1).

Als weitere Herausforderung bei der Erfassung handlungsnaher Aspekte professioneller Kompetenzen konnte im Verlauf der vorliegenden Arbeit identifiziert werden, dass die Operationalisierung des prozeduralen politikdidaktischen Wissens erweiterte Ansprüche an die Entwicklung geeigneter Testinstrumente stellt. Es wurde herausgearbeitet, dass dabei computerbasierten Erhebungsverfahren (Frey & Hartig, 2013; Klieme & Leutner, 2006, S. 882) und hier vor allem dem Format des video-cued testing (Blömeke, 2013b, S. 36), in dem Testpersonen auf Videovignetten mit anforderungsreichen und herausfordernden Unterrichtssituationen reagieren müssen, das Potenzial zukommt, handlungsnaher Aspekte professioneller Kompetenzen valider erfassen zu können als traditionelle schriftliche Tests (Lindmeier, 2013). Solche Möglichkeiten zur Erweiterung schriftlicher Leistungstests werden auch in der Politikdidaktik diskutiert (Gronostay et al., 2015, S. 162ff.; Weschenfelder, 2014a, S. 285), wobei entsprechende politikdidaktische Studien, die videobasierte Verfahren der Kompetenzdiagnostik fokussieren, derzeit noch nicht verfügbar sind. Mit dem Ziel, Möglichkeiten und Grenzen der Operationalisierung handlungsnaher Aspekte professioneller

politikdidaktischer Kompetenzen sowie der Implementation eines adäquaten computerbasierten Diagnoseinstrumentes diskutieren zu können (*Zweite Forschungsfrage*), wurde daher in der vorliegenden Arbeit eine explorative Studie durchgeführt, der sich an den Forschungen von Lindmeier (2011) bzw. Lindmeier et al. (2012) orientiert. Diese konnten für die Mathematik aufzeigen, dass es möglich ist, Unmittelbarkeit und Spontaneität durch den Einsatz von Videovignetten im computerbasierten Kompetenzdiagnostikprogramm vKid bei Probanden zu induzieren (Lindmeier, 2013, S. 54f.). Hierzu wurde in der vorliegenden Arbeit ein Leistungstest entwickelt, der das Fachwissen und das deklarative politikdidaktische Wissen – in Teilen traditionell und in Teilen computerbasiert – durch überwiegend geschlossene Items erfasst. Zur Diagnostik des prozeduralen politikdidaktischen Wissens wurden Videovignetten mit authentischen Anforderungssituationen aus unterschiedlichen videografierten Politikstunden entwickelt, die Verständnisschwierigkeiten, Fehler der Lernenden oder (methodische) Probleme im Unterrichtsverlauf enthalten und somit eine entsprechende Hilfestellung der Versuchspersonen erforderten. Die Probanden beantworteten die Items unter Zeitdruck mündlich in natürlicher Sprache in zwei unterschiedlichen Antwortformaten (Erklärung oder unterstützender Impuls). Diese offenen Antworten, aber auch die halb offenen Antworten aus der Skala zum deklarativen politikdidaktischen Wissen, wurden von zwei geschulten Ratern anhand taxonomischer Kategorien beurteilt und anschließend gemeinsam mit den weiteren Skalen quantitativ-statistisch analysiert.

Im Rahmen pilotierender Maßnahmen zeigte sich, dass die von Lindmeier (2011, S. 148ff.) entwickelten Skalen zur Beurteilung der offenen Antworten auch in der Politikdidaktik nutzbar sein können, wobei domänenspezifische Anpassungen notwendig waren, um die fachspezifische Qualität der Antworten beurteilen zu können. Hierzu wurden Merkmale hoher Unterrichtsqualität, die sich funktional auf das fachliche und fachdidaktische Wissen und Können von Lehrkräften beziehen, zunächst auf fächergenereller Ebene identifiziert und anschließend aus einer politikdidaktischen Perspektive reflektiert. Es wurde herausgearbeitet, dass eine systematische Auseinandersetzung mit Merkmalen hoher Unterrichtsqualität, die notwendigerweise auch fachspezifisch erfolgen muss (Klieme & Rakoczy, 2008, S. 229), in der Politikdidaktik noch nicht umfassend stattgefunden hat. Es konnten dennoch zentrale Aspekte der Qualität einer ‚guten‘ Rückmeldung im Politikunterricht probeweise charakterisiert werden; diese Aspekte wurden als kriteriale

Referenzpunkte im Ratingmanual zur Beurteilung der offenen Antworten in der Skala zum prozeduralen politikdidaktischen Wissen abgebildet. Die Beurteilerübereinstimmung auf Grundlage des Ratingmanuals war nach einer Nachschulung der Rater hoch. Da die Probanden einer Gelegenheitsstichprobe entstammen, die außerdem nicht alle Charakteristika einer Eichstichprobe erfüllt, sollte in zukünftigen Studien überprüft werden, inwieweit sich das Ratingmanual auch zur differenzierten Beurteilung der Leistungen von Probanden eignet, bei denen wesentlich höhere Merkmalsausprägungen in der Skala zum prozeduralen Wissen zu erwarten sind. Vor allem wäre zu begutachten, ob die Auflösung der in dieser Arbeit angewendeten 0/1/2-Skala Streuungen von Merkmalsausprägungen innerhalb der Niveaustufen ausreichend fein erfassen kann.

Die empirischen Befunde der in der vorliegenden Arbeit durchgeführten statistischen Verfahren wurden schließlich dazu genutzt, vorläufige Aussagen über die Zusammenhänge zwischen den Strukturkomponenten und damit über die zugrunde liegende Modellierung, über das Testinstrument sowie das Auswertungsverfahren und über differenzielle Aspekte der Leistungen von Probanden zu treffen (*Dritte Forschungsfrage*). Es zeigte sich zunächst, dass die Items der Skalen zum Fachwissen sowie zum deklarativen politikdidaktischen Wissen für die Testpersonen offenbar leichter zu beantworten waren als die Items der Skala zum prozeduralen politikdidaktischen Wissen. Dies entspricht den Annahmen, dass zum einen entsprechendes prozedurales Wissen bei den Lehramtsstudierenden wegen der lediglich begrenzten unterrichtspraktischen Erfahrungen, die sie im Studium (hier insbesondere in Praktika) sammeln können, eher gering ausgeprägt sein dürfte (Lindmeier, 2011, S. 187) und dass zum anderen das in den beiden weiteren Skalen erfasste deklarative Wissen eher einem ‚studiumsnahen‘ Wissen entspricht (Krauss et al., 2008b, S. 245; Neuweg, 2014, S. 591). Auch in der PKP-Studie erreichten Sekundarstufe-I-Studierende bessere Testleistungen in den Skalen zum Fachwissen und zum fachdidaktischen Wissen als die Referendarinnen und Referendare sowie Lehrkräfte (Weißeno et al., 2013, S. 192). Um hierzu genauere Aussagen treffen zu können, sollten überarbeitete Skalen der vorliegenden Studie zur Erhebung der politikdidaktischen Kompetenz von Referendarinnen und Referendare sowie Lehrkräften zum Einsatz kommen. Bedingt durch die Stichprobenziehung in der vorliegenden Studie ist zudem nicht auszuschließen, dass sich beispielsweise besonders leistungsstarke Studierende für eine Teilnahme gemeldet haben.

Die weiteren differenziellen Analysen zeigen, dass GyGe-Studierende auf statistisch signifikantem Niveau höhere Testleistungen erbracht haben als HRGe-Studierende; dieser Befund ist im Wesentlichen konsistent zu den PKP-Ergebnissen (ebd.). Gleichzeitig kann ein ebenfalls signifikanter Unterschied in den Testleistungen zwischen männlichen und weiblichen Probanden in allen Skalen festgestellt werden. Dies war wegen des nicht eindeutigen Bildes zu geschlechtsspezifischen Leistungen in anderen Studien nicht zwingend zu erwarten (Weschenfelder, 2014a, S. 271f.), zeigt sich aber wiederum im Kern konform zu den Befunden der PKP-Studie (ebd.). Die geschlechtsspezifischen Unterschiede könnten hierbei mit den Unterschieden zwischen HRGe und GyGe konfundiert sein. Probanden, die auch neben den obligatorischen Praktika, die im Lehramtsstudium zu absolvieren sind, bereits haupt- oder nebenberufliche Unterrichtserfahrungen gesammelt haben, zeigten insgesamt keine höheren Testleistungen als Studierende ohne entsprechende Unterrichtserfahrungen. Wegen der deutlich unterschiedlichen Gruppengröße ist dieser Befund jedoch entsprechend vorsichtig zu interpretieren. Auch ein – relativ zu den durchschnittlichen Unterrichtserfahrungen in Zeitstunden – größerer zeitlicher Umfang an eigenen Lehrerfahrungen zeigte keine statistisch signifikanten Unterschiede zu relativ gering ausgeprägten eigenen Unterrichtserfahrungen, wobei die durchschnittliche Testleistung der ‚erfahreneren‘ Studierenden leicht über den ‚unerfahrenen‘ Studierenden liegt. Zwar ist dieser Befund nicht erwartungskonform, da theoretisch von besseren Leistungen der Studierenden mit mehr Erfahrungswissen in der Skala zum prozeduralen politikdidaktischen Wissen ausgegangen werden kann, jedoch ist zu berücksichtigen, dass die Differenz zwischen den Merkmalsausprägungen (Unterrichtserfahrung hoch versus niedrig in Zeitstunden) eher gering ausgeprägt ist, insbesondere wenn die Streuung der Werte beachtet wird. Mehr Aufschluss hierüber könnte in zukünftigen Studien der Vergleich mit Kontrastgruppen wie bspw. Lehrkräften mit langjähriger Berufserfahrung geben (vgl. auch Krauss et al., 2008a, S. 880ff.).

Die empirischen Befunde zur Itemtrennschärfe sowie zu den Korrelationen zwischen den einzelnen Strukturkomponenten weisen darauf hin, dass es sich bei diesen um voneinander unterscheidbare Wissenskomponenten handelt, die gleichzeitig aber zusammenhängen. Diese vorläufige Erkenntnis ist anschlussfähig an die Befunde der bereits angesprochenen Studien zur professionellen Kompetenz von Lehrkräften (Kunter et al., 2011; Lindmeier, 2011; Schmelzing, 2010; Weißeno et al., 2013; Weschenfelder, 2014a). Der Umstand, dass die

interne Konsistenz der Skala zum Fachwissen nicht sinkt, wenn probeweise der Reliabilitätskoeffizient für eine gemeinsame Skala berechnet wird, deutet darauf hin, dass die Items etwas Ähnliches messen, was insofern mit den bereits angesprochenen theoretischen Annahmen einer deutlichen Verbindung des Fachwissens mit dem fachdidaktischen Wissen vereinbar ist (vgl. Kapitel 2.2). Bevor dieser Befund der vorliegenden Studie jedoch als Indikator für eine „implizite fachdidaktische Dimension des Fachwissens“ (Neuweg, 2014, S. 591) angesehen werden kann, sind weitere Untersuchungen hierzu notwendig, denn hier könnte sich durchaus auch ein Effekt des höheren Auflösungsgrades der Modellierung sowie der inhaltspezifischen ‚Füllung‘ der einzelnen Wissenskomponenten abzeichnen: Es erscheint nachvollziehbar, dass die Beantwortung von lernprozessbezogenen Fragen zu einem bestimmten Inhaltsbereich in der Skala zum deklarativen politikdidaktischen Wissen entsprechende fachwissenschaftliche Kenntnisse voraussetzt und sich die Testleistungen in beiden Skalen daher auf einem ähnlichen Niveau bewegen (vgl. auch Riese & Reinhold, 2010, S. 182). Aus theoretischer Perspektive (Neuweg, 2014), aber auch wegen der Schwierigkeit, Test-Items zu entwickeln, die exklusiv das fachdidaktische Wissen erfassen sollen (Krauss et al., 2008a, S. 888), ist bei der Item-(Weiter-)Entwicklung auf diesen Aspekt ein besonderes Gewicht zu legen. Um weitere differenzielle Befunde zu den Zusammenhängen zwischen dem Fachwissen und dem fachdidaktischen Wissen im Sinne der Modellierung in dieser Arbeit erlangen zu können, wäre in zukünftigen Forschungen der Fokus auch auf mehrere Inhaltsbereiche zu richten. Hierdurch wären schließlich erst detailliertere Aussagen zur Robustheit – verstanden als „Ausmaß der Invarianz von aus dem Modell abgeleiteten Aussagen und Parametern von ... Variationen von Operationalisierungen (Items)“ (Robitzsch, 2013, S. 42) – der Modellierung möglich. Durch die höhere Feinkörnigkeit der Messung in dieser Dimension (aber auch in der Dimension der Wissensarten) kann eine stärkere Sensitivität gegenüber entsprechenden Variationen angenommen werden (ebd.).⁶²

Die Skala zum prozeduralen politikdidaktischen Wissen lässt sich schließlich deutlicher von den beiden Skalen zum deklarativen Wissen unterscheiden, was die Annahme stützt, dass zur Beantwortung der Items tatsächlich ein anderes Wissen vonnöten ist und dass das Vorhandensein adäquaten Fachwissens sowie deklarativen politikdidaktischen Wissens nicht

⁶² Dass Kompetenzstrukturmodelle „nicht situationsinvariant“ (Robitzsch, 2013, S. 44) sind und insofern in diesem Aspekt weniger robust sind, ist gleichsam in der Kontextspezifität von Kompetenzen begründet (ebd.).

ausreicht, um in dieser Skala Antworten auf hohem Niveau geben zu können. Unbeantwortet muss zunächst bleiben, woher sich das prozedurale politikdidaktische Wissen speist, da die differenziellen Befunde aufgrund der Stichprobencharakteristika und der Bodeneffekte der Testleistungen in dieser Skala vorsichtig interpretiert werden müssen. Lindmeier et al. (2012) gehen davon aus, dass das in ihrer Studie konzeptualisierte Konstrukt der aktionsbezogenen Kompetenz gemessen werden kann, „wenn praktische Erfahrung vorliegt“ (S. 115), weshalb es in künftigen Forschungen daher aufschlussreich wäre, wie beispielsweise Politiklehrkräfte mit langjähriger Berufserfahrung auf Items in der Skala zum prozeduralen politikdidaktischen Wissen antworten. Erst dadurch wäre auch eine Normierung der Items der Skala – ebenso wie der Items der übrigen Skalen – möglich. Hierzu scheint auch eine Erweiterung des Gesamttests um Skalen zu professionellen Überzeugungen der (angehenden) Lehrkräfte sinnvoll zu sein, um genauer analysieren zu können, welche Aspekte professioneller Kompetenzen als handlungsleitend identifiziert werden können.

8.2 Ausblick

Wegen des erkundenden Charakters der vorliegenden Studie sowie vor dem Hintergrund der notwendigen weiteren (Teil-)Entwicklungsschritte zum Zwecke der Nutzung von Kompetenzmodellen (vgl. auch Abbildung 10) sind mögliche Implikationen, die sich für zukünftige Forschungen und – mehr noch – für die Lehrerbildung ergeben können, bedachtsam zu formulieren. Die theoretische Modellexploration und die empirischen Befunde der explorativen Studie lassen jedoch den Schluss zu, dass es prinzipiell möglich ist, handlungsnahen Aspekte professioneller politikdidaktischer Kompetenzen in einem erweiterten mehrdimensionalen Strukturmodell abzubilden, zu operationalisieren und im Rahmen eines computerbasierten Erhebungsverfahrens (video-cued testing) zu erfassen. Eine Anschlussfähigkeit der Modellierung an das vorliegende Kompetenzstrukturmodell der PKP-Studie ist dabei gegeben. In zukünftigen Studien sollte dabei das Erhebungsinstrument verbessert und an einer Stichprobe eingesetzt werden, die neben Studierenden auch Referendarinnen und Referendare sowie Politiklehrkräfte umfasst. Hierdurch könnten noch aussagekräftigere differenzielle Befunde zu Extremgruppen generiert werden (zum Beispiel *keine Unterrichtserfahrungen versus langjährige Berufserfahrungen*) und die Strukturkomponenten des Modells könnten noch genauer analysiert werden. Die Exklusivität

der konzeptualisierten Wissenskomponenten sollte durch den Vergleich mit Kontrastpopulationen empirisch abgesichert werden (z. B. durch den Vergleich des politikdidaktischen Wissens und Könnens von erfahrenen Lehrkräften und Fachwissenschaftlern) (Schmelzing, 2010, S. 118). Ein methodisches Vorgehen auf Basis der IRT ist dabei angezeigt, wozu auch eine größere Stichprobenziehung notwendig ist.

Schließlich konnten aufgrund des Fokus der vorliegenden Arbeit weitere wichtige Aspekte der professionellen Handlungskompetenz wie das pädagogisch-psychologische Wissen, professionelle Überzeugungen (Weschenfelder, 2014a) oder auch reflexive Kompetenzen (Lindmeier, 2011; Lindmeier et al., 2012; Schmelzing, 2010) nicht berücksichtigt werden. Das in dieser Studie erprobte Erhebungsverfahren ermöglicht dabei durchaus die Implementation bestehender Forschungsstrategien. So wäre es vorstellbar, auch handlungsnahe Aspekte pädagogisch-psychologischen Wissens videobasiert zu erfassen oder aber Videostimuli als Ausgangspunkt für den Aufbau und die Diagnostik reflexiver Kompetenzen zu nutzen (vgl. z. B. Sherin & van Es, 2009).

Voraussetzung hierfür ist der Einsatz eines computerbasierten Diagnoseverfahrens, wie in dieser Studie in Form des Computerprogramms vKid. Erwartungsgemäß war der technische Aufwand in der Entwicklung des Tests durch die zusätzlichen Arbeitsschritte wie Videovignettenentwicklung, technische Erprobung, Anschaffung weiterer technischer Peripheriegeräte hoch,⁶³ jedoch ist hierbei zu bedenken, dass ein einmalig validiertes computerbasiertes Erhebungsverfahren in Folgeerhebungen (relativ) kostenneutral eingesetzt werden kann (Frey & Hartig, 2013, S. 56). Es zeigte sich zudem in dieser Erhebung, dass das optional zu verwendende videografierte Notizklemmbrett (vgl. Kapitel 6.2) von lediglich etwa einem Drittel aller Probanden genutzt wurde. Die Notizen wurden nicht in die Datenauswertung mit einbezogen, da ihr Informationsgehalt gering war. Wenn in einer Folgestudie daher auf die Videografie des Notizklemmbretts verzichtet würde, wären die einmaligen Anschaffungskosten für ein Headset pro Computer relativ niedrig. Die sehr geringe Anzahl an technisch bedingt fehlenden Werten bestätigt die hohe Systemstabilität des Computerprogramms.

Computerbasierte Diagnoseverfahren könnten in Verbindung mit dem Einsatz von Videovignetten auch für das universitäre Lehramtsstudium und die Lehrerfortbildung neue Impulse

⁶³ Hierunter gefasst ist noch nicht die eigentliche Videografie der Unterrichtsstunden, die noch weitere Arbeitsschritte umfasst (Petko et al., 2003, S. 268ff.).

liefern, da „[d]erzeitige Studien zur Kompetenzdiagnostik ... ihr Potential als Unterstützungs- und Feedbackinstrument für Lehrkräfte bislang noch kaum aus[schöpfen]“ (Bögeholz & Eggert, 2013, S. 61). In der Annahme, dass auch spezialisierte Kompetenzen durch zielgerichtete Übungen aufgebaut werden können, wäre es beispielsweise denkbar, den Ansatz des video-cued testing zu Übungszwecken in Seminaren zu nutzen. So könnten in einem ersten Schritt eigene oder fremde Unterrichtsvideos anhand *structured viewing guides* von Studierenden analysiert werden (Brouwer, Besselink & Muller, 2011), um hierdurch reflexive Kompetenzen aufzubauen (Newhouse, Lane & Brown, 2007). Anschließend könnten inhaltlich passende Items mit Videovignetten, auf die die Studierenden unter Zeitdruck spontan reagieren müssen, dazu dienen, handlungsnahes Wissen aufzubauen. In ähnlichen Lehr-Lern-Settings wäre es auch vorstellbar, diagnostische Kompetenzen von (angehenden) Lehrkräften zu (Fehl-)Vorstellungen der Lernenden aufzubauen (Bögeholz & Eggert, 2013, S. 61; Pant, 2013, S. 75).

Auf einer systemischen Ebene wäre – ein entsprechend validiertes Testinstrument vorausgesetzt – das Erhebungsverfahren des video-cued testing auch dazu geeignet, Kompetenzzuwächse im Lehramtsstudium zu diagnostizieren und insofern „politisches oder bildungsadministratives Steuerungswissen zu generieren“ (Pant, 2013, S. 73). So wäre beispielsweise zu erwarten, dass sich das Absolvieren eines umfangreichen Praktikums – wie dem Praxissemester in Nordrhein-Westfalen – mit entsprechenden unterrichtspraktischen Bestandteilen positiv auf den Aufbau des prozeduralen politikdidaktischen Wissens auswirkt. Solche angenommenen Effekte könnten mit einem adäquaten computerbasierten Verfahren erfasst werden, wodurch sich erweiterte Möglichkeiten zur Evaluation von neuen Ausbildungsstrategien ergeben können.

Die politikdidaktische Kompetenzforschung kann noch nicht auf eine jahrelange und umfassende Forschungstradition zurückblicken. Dies trifft insbesondere auf die Modellierung und Operationalisierung handlungsnaher Aspekte professioneller Kompetenzen von Politiklehrkräften zu. Verbunden mit der Erkenntnis, dass Lehrerinnen und Lehrer maßgeblich zum Lernerfolg der Schülerinnen und Schüler beitragen, ergibt sich für die Politikdidaktik die Herausforderung, domänenspezifische Kompetenzen diagnostizieren zu können. Ansätze und Möglichkeiten hierzu für das politikdidaktische Wissen und Können sollte die vorliegende Arbeit aufzeigen.

9 Literaturverzeichnis

- Alisch, L.-M., Hermkes, R. & Möbius, K. (2009). Messen von Lehrprofessionalität I: Grundlagen. In E. Terhart, H. Bennewitz & M. Rothland (Hrsg.), *Handbuch der Forschung zum Lehrerberuf* (2. Aufl., S. 249-262). Münster: Waxmann.
- Amelang, M. & Schmidt-Atzert, L. (2006). *Psychologische Diagnostik und Intervention* (4. Aufl.). Berlin: Springer Medizin.
- Anderson, J. R. (1983). *The architecture of cognition*. Cambridge: Harvard University Press.
- Anderson, J. R. (2013). *Kognitive Psychologie* (7. Aufl.). Berlin: Springer VS.
- Autorengruppe Fachdidaktik (2011). Sozialwissenschaftliche Basiskonzepte als Leitideen der politischen Bildung – Perspektiven für Wissenschaft und Praxis. In Autorengruppe Fachdidaktik (Hrsg.), *Konzepte der politischen Bildung. Eine Streitschrift* (Reihe Politik und Bildung, Bd. 64, S. 163-171). Schwalbach/Ts.: Wochenschau.
- Baumert, J. & Köller, O. (2000). Unterrichtsgestaltung, verständnisvolles Lernen und multiple Zielerreichung im Mathematik- und Physikunterricht der gymnasialen Oberstufe. In J. Baumert, W. Bos & R. Lehmann (Hrsg.), *TIMSS/III. Dritte Internationale Mathematik- und Naturwissenschaftsstudie – Mathematische und naturwissenschaftliche Bildung am Ende der Schullaufbahn* (Bd. 2: Mathematische und physikalische Kompetenzen in der Oberstufe, S. 271-315). Opladen: Leske + Budrich.
- Baumert, J. & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. *Zeitschrift für Erziehungswissenschaften*, 4, 469-520.
- Baumert, J. & Kunter, M. (2011a). Das Kompetenzmodell von COACTIV. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV* (S. 29-53). Münster: Waxmann.
- Baumert, J. & Kunter, M. (2011b). Das mathematikspezifische Wissen von Lehrkräften, kognitive Aktivierung im Unterricht und Lernfortschritte von Schülerinnen und Schülern. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV* (S. 163-192). Münster: Waxmann.

- Baumert, J., Kunter, M., Blum, W., Klusmann, U., Krauss, S. & Neubrand, M. (2011). Professionelle Kompetenz von Lehrkräften, kognitiv aktivierender Unterricht und die mathematische Kompetenz von Schülerinnen und Schülern (COACTIV) – Ein Forschungsprogramm. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV* (S. 7-25). Münster: Waxmann.
- Besser, M. (2014). *Lehrerprofessionalität und die Qualität von Mathematikunterricht. Quantitative Studien zu Expertise und Überzeugungen von Mathematiklehrkräften* (Perspektiven der Mathematikdidaktik). Wiesbaden: Springer Fachmedien.
- Besser, M. & Krauss, S. (2009). Zur Professionalität als Expertise. In O. Zlatkin-Troitschanskaia, K. Beck, D. Sembill, R. Nickolaus & R. Mulder (Hrsg.), *Lehrerprofessionalität. Bedingungen, Genese, Wirkungen und ihre Messung* (S. 71-82). Weinheim: Beltz.
- Blömeke, S. (2002). Universität und Lehrerbildung. Bad Heilbrunn: Klinkhardt.
- Blömeke, S. (2013a). Einleitung: Professionelle Kompetenzen im Studienverlauf. In S. Blömeke, A. Bremerich-Vos, G. Kaiser, G. Nold, H. Haudeck, J.-U. Keßler & K. Schwippert (Hrsg.), *Professionelle Kompetenzen im Studienverlauf. Weitere Ergebnisse zur Deutsch-, Englisch- und Mathematiklehrerbildung aus TEDS-LT* (S. 7-24). Münster: Waxmann.
- Blömeke, S. (2013b). Moving to a higher state of confusion. Der Beitrag der Videoforschung zur Kompetenzforschung. In U. Riegel & K. Macha (Hrsg.), *Videobasierte Kompetenzforschung in den Fachdidaktiken* (Fachdidaktische Forschungen, Bd. 4, S. 25-43). Münster: Waxmann.
- Blömeke, S., Bremerich-Vos, A., Haudeck, H., Kaiser, G., Nold, G., Schwippert, K. & Willenberg, H. (Hrsg.). (2011). *Kompetenzen von Lehramtsstudierenden in gering strukturierten Domänen. Erste Ergebnisse aus TEDS-LT*. Münster: Waxmann.
- Blömeke, S., Bremerich-Vos, A., Kaiser, G., Nold, G., Haudeck, H., Keßler, J.-U. & Schwippert, K. (Hrsg.) (2013). *Professionelle Kompetenzen im Studienverlauf. Weitere Ergebnisse zur Deutsch-, Englisch- und Mathematiklehrerbildung aus TEDS-LT*. Münster: Waxmann.
- Blömeke, S., Kaiser, G. & Lehmann, R. (Hrsg.). (2008). *Professionelle Kompetenz angehender Lehrerinnen und Lehrer. Wissen, Überzeugungen und Lerngelegenheiten*

- angehender Mathematikstudierender und -referendare. Erste Ergebnisse zur Wirksamkeit der Lehrerbildung. Münster: Waxmann.
- Blömeke, S., Kaiser, G. & Lehmann, R. (Hrsg.). (2010a). *TEDS-M 2008. Professionelle Kompetenz und Lerngelegenheiten angehender Mathematiklehrkräfte für die Sekundarstufe I im internationalen Vergleich*. Münster: Waxmann.
- Blömeke, S., Kaiser, G. & Lehmann, R. (Hrsg.). (2010b). *TEDS-M 2008. Professionelle Kompetenz und Lerngelegenheiten angehender Primarstufenlehrkräfte im internationalen Vergleich*. Münster: Waxmann.
- Blömeke, S., König, J., Busse, A., Suhl, U., Benthien, J., Döhrmann, M. et al. (2014). Von der Lehrerbildung in den Beruf – Fachbezogenes Wissen als Voraussetzung für Wahrnehmung, Interpretation und Handeln im Unterricht. *Zeitschrift für Erziehungswissenschaft*, 17, 509-542.
- BMBF (Hrsg.) (2001). *TIMSS – Impulse für Schule und Unterricht. Forschungsbefunde, Reforminitiativen, Praxisberichte und Video-Dokumente*. Bonn: Herausgeber. Abgerufen am 27.03.2015 unter <http://www.ssg-bildung.ub.uni-erlangen.de/timss.pdf>
- Bögeholz, S. & Eggert, S. (2013). Welche Rolle spielt Kompetenzdiagnostik im Rahmen von Lehr-Lernprozessen? *Zeitschrift für Erziehungswissenschaft*, 16(1), 59-64.
- Bortz, J. & Döring, N. (2006). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler* (4. Aufl.). Heidelberg: Springer Medizin.
- Bortz, J. & Lienert, G. A. (2008). *Kurzgefasste Statistik für die Klinische Forschung. Leitfaden für die verteilungsfreie Analyse kleiner Stichproben* (3. Aufl.). Heidelberg: Springer Medizin.
- Bortz, J., Lienert, G. A. & Boehnke, K. (2008). *Verteilungsfreie Methoden in der Biostatistik* (3. Aufl.). Heidelberg: Springer Medizin.
- Bowen, D. J., Kreuter, M., Spring, B., Cofta-Woerpel, L., Linnan, L., Weiner, D. et al. (2009). How we design feasibility studies. *American Journal of Preventive Medicine*, 36(5), 452-457.
- Bromme, R. (1992, 2014). *Der Lehrer als Experte. Zur Psychologie des professionellen Wissen* (Standardwerke aus Psychologie und Pädagogik, Bd. 7) Münster: Waxmann.
- Brouwer, N., Besselink, E. & Muller, G. (2011, June). *Self-viewing with structured viewing guides. First results of an intervention study*. Paper presented at the International Symposium Research on the Use of Video in the Professional Development of Mathematics and Science Teachers, Lausanne, Switzerland. Abgerufen am

- 12.03.2015 unter http://education.uci.edu/docs/Self-viewing%20with%20structured%20viewing%20guides_Brouwer%20et%20al_June2011.pdf
- Brovelli, D., Bölsterli, K., Rehm, M. & Wilhelm, M. (2013). Erfassen professioneller Kompetenzen für den naturwissenschaftlichen Unterricht: Ein Vignettentest mit authentisch komplexen Unterrichtssituationen und offenem Antwortformat. *Unterrichtswissenschaft*, 41(4), 306-329.
- Brunner, M., Anders, Y., Hachfeld, A. & Krauss, S. (2011). In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV* (S. 215-234). Münster: Waxmann.
- Bühner, M. (2006). *Einführung in die Test- und Fragebogenkonstruktion* (2. Aufl.). München: Pearson.
- Bühner, M. & Ziegler, M. (2009). *Statistik für Psychologen und Sozialwissenschaftler*. München: Pearson.
- Carey, S. (1999). Knowledge acquisition: Enrichment or conceptual change? In E. Margolis & S. Laurence (Hrsg.), *Concepts: Core readings* (S. 459-487). Cambridge: MIT Press.
- Clement, J. (2008). The role of explanatory models in teaching for conceptual change. In S. Vosniadou (Hrsg.), *International handbook of research on conceptual change* (S. 417-452). New York: Routledge.
- De Jong, T. & Ferguson-Hessler, M. G. M. (1996). Types and qualities of knowledge. *Educational Psychologist*, 31(2), 105-113. Abgerufen am 16.01.2015 unter <http://doc.utwente.nl/26717/1/types.pdf>
- Deppermann, A. (2008). *Gespräche analysieren* (Qualitative Sozialforschung, Bd. 3, 4. Aufl.). Opladen: Leske + Budrich.
- Detjen, J. (2006). Wie viel Wirtschaft braucht die politische Bildung? In G. Weißenö (Hrsg.), *Politik und Wirtschaft unterrichten* (S. 62-79). Wiesbaden: VS.
- Detjen, J. (2007). *Politische Bildung*. München: Oldenbourg.
- Detjen, J., Massing, P., Richter, D. & Weißenö, G. (2012). *Politikkompetenz – ein Modell*. Wiesbaden: Springer VS.

- Dübbelde, G. (2013). *Diagnostische Kompetenzen angehender Biologie-Lehrkräfte im Bereich naturwissenschaftliche Erkenntnisgewinnung*. Dissertation, Universität Kassel.
- Dubberke, T., Kunter, M., McElvany, N., Brunner, M. & Baumert, J. (2008). Lerntheoretische Überzeugungen von Mathematiklehrkräften. Einflüsse auf die Unterrichtsgestaltung und den Lernerfolg von Schülerinnen und Schülern. *Zeitschrift für Pädagogische Psychologie*, 22(3/4), 193-206.
- DVPW (2003). *Gestufte Studiengänge in der Politikwissenschaft – Empfehlungen zu einem Kerncurriculum von Vorstand und Beirat der DVPW*. Abgerufen am 14.03.2015 unter <http://www.dvpw.de/fileadmin/docs/Kerncurriculum.pdf>
- Eraut, M. (1994/2000). *Developing professional knowledge and competence*. London: Routledge/Falmer.
- Eraut, M. (1998). Concepts of competence. *Journal of Interprofessional Care*, 12(2), 127-139.
- Ericsson, K. A., Krampe, R. T. & Tesch-Romer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, 100(3), 363-406.
- Faissner, E., Flöry, P., Jaklin, J., Kiss, A., Koliander, B., Lang, O. et al. (o. J.). Naturwissenschaftliche Bildungsstandards. Berufsbildende Höhere Schulen. Das Kompetenzmodell. Abgerufen am 06.09.2015 unter http://www.berufsbildendeschulen.at/fileadmin/content/bbs/AGBroschueren/NaturwissenschaftenBHS_Vers.09.pdf
- Fankhauser, R. (2013). Videobasierte Unterrichtsbeobachtung: die Quadratur des Zirkels? *Forum Qualitative Sozialforschung*, 14(1), Art. 24. Abgerufen am 14.12.2014 unter <http://nbn-resolving.de/urn:nbn:de:0114-fqs1301241>
- Fend, H. (2008) *Neue Theorie der Schule. Einführung in das Verstehen von Bildungssystemen* (2. Aufl.). Wiesbaden: VS.
- Fenstermacher, G. D. (1994). The knower and the known: The nature of knowledge in research on teaching. In L. Darling-Hammond (Hrsg.), *Review of research education* (Bd. 20, S. 3-56). Washington: American Educational Research Association.
- Festinger, L (2012). *Theorie der kognitiven Dissonanz* (2. Aufl.). Bern: Huber.
- Fischer, K. W., Bullock, D., Rotenberg, E. & Raya, P. (1993). The dynamics of competence: How context contributes directly to skill. In R. H. Wozniak & K. W. Fischer (Hrsg.), *Development in context: Acting and thinking in specific environments* (S. 93-117).

- Hillsdale: Erlbaum Associates. Abgerufen am 16.05.2014 unter <https://www.gse.harvard.edu/~ddl/articlesCopy/FischerBullockRotenbergdynofcomp1993.pdf>
- Fleischer, J., Koeppen, K., Kenk, M., Klieme, E. & Leutner, D. (2013). Kompetenzmodellierung: Struktur, Konzepte und Forschungszugänge des DFG-Schwerpunktprogramms. *Zeitschrift für Erziehungswissenschaft*, 16, 5-22.
- Frey, A. (2014). Kompetenzmodelle und Standards in der Lehrerbildung und im Lehrerberuf. In E. Terhart, H. Bennewitz & M. Rothland (Hrsg.), *Handbuch der Forschung zum Lehrerberuf* (2. Aufl., S. 712-744). Münster: Waxmann.
- Frey, A. & Hartig, J. (2013). Wann sollten computerbasierte Verfahren zur Messung von Kompetenzen anstelle von papier- und bleistift-basierten Verfahren eingesetzt werden? *Zeitschrift für Erziehungswissenschaft*, 16, 53-57.
- GFD (2004). *Kerncurriculum Fachdidaktik. Orientierungsrahmen für alle Fachdidaktiken*. Abgerufen am 16.04.2015 unter http://fachdidaktik.org/cms/download.php?cat=Ver%C3%B6ffentlichungen&file=Publikationen_zur_Lehrerbildung-Anlage_3.pdf
- GFD (2005). *Fachdidaktische Kompetenzbereiche, Kompetenzen und Standards für die 1. Phase der Lehrerbildung (BA+MA)*. Abgerufen am 16.04.2015 unter http://fachdidaktik.org/cms/download.php?cat=Ver%C3%B6ffentlichungen&file=Publikationen_zur_Lehrerbildung-Anlage_1.pdf
- Gess-Newsome, J. (1999). Pedagogical content knowledge: An introduction and orientation. In J. Gess-Newsome & N. G. Lederman, *Examining pedagogical content knowledge* (S. 3-17). Dordrecht: Kluwer.
- Girmes, R. (2006). Lehrerprofessionalität in einer demokratischen Gesellschaft. Über Kompetenzen und Standards in einer erziehungswissenschaftlich fundierten Lehrerbildung. In C. Allemann-Ghionda & E. Terhart (Hrsg.), *Kompetenzen und Kompetenzentwicklung von Lehrerinnen und Lehrern* (Zeitschrift für Pädagogik, Beiheft 51, S. 14-29). Weinheim: Beltz.
- Götzmann, A. (2008). Politische Konzepte von Grundschüler/-innen zu Öffentlichkeit. In G. Weißenö (Hrsg.), *Politikkompetenz. Was Unterricht zu leisten hat* (Schriftenreihe Bd. 645, S. 293-308). Bonn: bpb.

- Grammes, T. (2005). Kontroversität. In W. Sander (Hrsg.), *Handbuch politische Bildung* (Reihe Politik und Bildung, Bd. 32, 3. Aufl., S. 126-145). Schwalbach/Ts.: Wochenschau.
- Gramzow, Y., Riese, J. & Reinhold, P. (2013). Modellierung fachdidaktischen Wissens angehender Physiklehrkräfte. *Zeitschrift für Didaktik der Naturwissenschaften*, 19, 7-30.
- Gramzow, Y., Riese, J. & Reinhold, P. (2013). Modellierung fachdidaktischen Wissens angehender Physiklehrkräfte. *Zeitschrift für Didaktik der Naturwissenschaften*, 19, 7-30.
- Gronostay, D. (2015). *Dokumentation der Videostudie „Argumentative Lehr-Lernprozesse“: Erhebungsinstrumente, Unterrichtsmaterialien und Datenaufbereitung*. Unveröffentlichtes Manuskript, Universität Duisburg-Essen.
- Gronostay, D., Neumann, D. & Manzel, S. (2015). Videographie in der politikdidaktischen Forschung – Reichweiten und Grenzen anhand aktueller Forschungsprojekte. In A. Petrik (Hrsg.), *Formate fachdidaktischer Forschung in der politischen Bildung* (Schriftenreihe der Gesellschaft für Politikdidaktik und politische Jugend- und Erwachsenenbildung, Bd. 14, S. 158-166). Schwalbach/Ts.: Wochenschau.
- Hahn-Laudenberg, K., Weißeno, G. & Eck, V. (2010). Wie verknüpfen Schüler/-innen Fachkonzepte in Lernaufgaben? Eine computergestützte Analyse von Schülerantworten aus der TEESAEC-Studie. In I. Juchler (Hrsg.), *Kompetenzen in der politischen Bildung* (Schriftenreihe der Gesellschaft für Politikdidaktik und politische Jugend- und Erwachsenenbildung, Bd. 9, S. 159-168). Schwalbach/Ts.: Wochenschau.
- Hair, J. F. (Jr.), Black, W. C. Babin, B. J. & Anderson, R. E. (2010). *Multivariate data analysis. A global perspective* (7. Aufl.). Upper Saddle River: Pearson.
- Hallet, W. (2006). *Didaktische Kompetenzen. Lehr- und Lernprozesse erfolgreich gestalten*. Stuttgart: Klett.
- Hameyer, U. & Röber, B. (2011). Entdeckendes Lernen. In J. Wiechmann (Hrsg.), *Zwölf Unterrichtsmethoden. Vielfalt für die Praxis* (5. Aufl., S. 129-143). Weinheim: Beltz.
- Hartig, J. (2009). Messung der Kompetenzen von Lehrpersonen mit Modellen der Item-Response-Theorie. In O. Zlatkin-Troitschanskaia, K. Beck, D. Sembill, R. Nickolaus & R. Mulder (Hrsg.), *Lehrerprofessionalität. Bedingungen, Genese, Wirkungen und ihre Messung* (S. 295-310). Weinheim: Beltz.

- Hartig, J. & Frey, A. (2013). Sind Modelle der Item-Response-Theorie (IRT) das „Mittel der Wahl“ für die Modellierung von Kompetenzen? *Zeitschrift für Erziehungswissenschaft*, 16, 47-51.
- Hartig, J. & Höhler, J. (2010). Modellierung von Kompetenzen mit mehrdimensionalen IRT-Modellen. Projekt MIRT. In E. Klieme, D. Leutner & M. Kenk (Hrsg.), *Kompetenzmodellierung. Zwischenbilanz des DFG-Schwerpunktprogramms und Perspektiven des Forschungsansatzes* (Zeitschrift für Pädagogik, Beiheft 56, S. 189-198). Weinheim: Beltz.
- Hartig, J. & Jude, N. (2007). Empirische Erfassung von Kompetenzen und psychometrische Kompetenzmodelle. In J. Hartig & E. Klieme (Hrsg.), *Möglichkeiten und Voraussetzungen technologiebasierter Kompetenzdiagnostik. Eine Expertise im Auftrag des Bundesministeriums für Bildung und Forschung* (Bildungsforschung Bd. 20, S. 17-36). Berlin: BMBF.
- Hartig, J. & Klieme, E. (2006). Kompetenz und Kompetenzdiagnostik. In K. Schweizer (Hrsg.), *Leistung und Leistungsdiagnostik* (S. 127-143). Heidelberg: Springer Medizin.
- Hascher, T. (2014). Forschung zur Wirksamkeit der Lehrerbildung. In E. Terhart, H. Bennewitz & M. Rothland (Hrsg.), *Handbuch der Forschung zum Lehrerberuf* (2. Aufl., S. 542-571). Münster: Waxmann.
- Hattie, J. (2009). *Visible learning. A synthesis of over 800 meta-analyses relating to achievement*. London: Routledge.
- Hedtke, R. (2002). Die Kontroversität in der Wirtschaftsdidaktik. *Gesellschaft, Wirtschaft, Politik*, 51(2), 173-186. Abgerufen am 17.04.2015 unter <http://www.uni-bielefeld.de/soz/ag/hedtke/pdf/kontroversitaet.pdf>
- Hedtke, R. (2006). Integrative politische Bildung. Sozialwissenschaftliche Kompetenzen jenseits von Fächern und Disziplinen. In G. Weißeno, *Politik und Wirtschaft unterrichten* (S. 216-229). Wiesbaden: VS.
- Hedtke, R. (2011). Die politische Domäne im sozialwissenschaftlichen Feld. In Autorengruppe Fachdidaktik (Hrsg.), *Konzepte der politischen Bildung. Eine Streitschrift* (Reihe Politik und Bildung, Bd. 64, S. 51-68). Schwalbach/Ts.: Wochenschau.
- Helmke, A. (2014). *Unterrichtsqualität und Lehrerprofessionalität. Diagnose, Evaluation und Verbesserung des Unterrichts* (5. Aufl.). Seelze: Klett-Kallmeyer.

- Helmke, A., Helmke, T. & Schrader, F.-W. (2007). Unterrichtsqualität: Brennpunkte und Perspektiven der Forschung. In K.-H. Arnold (Hrsg.), *Unterrichtsqualität und Fachdidaktik* (S. 51-72). Bad Heilbrunn: Julius Klinkhardt.
- Helsper, W. (2014). Lehrerprofessionalität – der strukturtheoretische Professionsansatz zum Lehrerberuf. In E. Terhart, H. Bennewitz & M. Rothland (Hrsg.), *Handbuch der Forschung zum Lehrerberuf* (2. Aufl., S. 216-240). Münster: Waxmann.
- Hill, H. C., Rowan, B. & Loewenberg Ball, D. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American Educational Research Journal*, 42(2), 371-406.
- Hill, H. C., Schilling, S. G. & Loewenberg Ball, D. (2004). Developing measures of teachers' mathematics knowledge for teaching. *The Elementary School Journal*, 105(1), 11-30. Abgerufen am 14.03.2015 unter <http://sitemaker.umich.edu/lmt/files/hillshillingball.pdf>
- Hofer, R. (2012). *Wissen und Können. Begriffsanalytische Studien zu einer kompetenzorientierten Wissensbildung am Gymnasium* (Internationale Hochschulschriften, Bd. 561). Münster: Waxmann.
- Hufer, K.-P. & Richter, D. (2013). Politische Bildung als Profession. Verständnisse und Forschungen. In K.-P. Hufer & D. Richter (Hrsg.), *Politische Bildung als Profession. Verständnisse und Forschungen* (Perspektiven politischer Bildung, Schriftenreihe Bd. 1355, S. 13-19). Bonn: bpb.
- Huppert, A. & Abs, H. J. (2013). Profession, Professionalisierung und Professionalität im Lehrerberuf. In K.-P. Hufer & D. Richter (Hrsg.), *Politische Bildung als Profession. Verständnisse und Forschungen* (Perspektiven politischer Bildung, Schriftenreihe Bd. 1355, S. 65-80). Bonn: bpb.
- Hussy, W., Schreier, M. & Echterhoff, G. (2013). *Forschungsmethoden in Psychologie und Sozialwissenschaften für Bachelor* (2. Aufl.). Berlin: Springer.
- Jäger, R. S. (2009). In O. Zlatkin-Troitschanskaia, K. Beck, D. Sembill, R. Nickolaus & R. Mulder (Hrsg.), *Lehrerprofessionalität. Bedingungen, Genese, Wirkungen und ihre Messung* (S. 106-116). Weinheim: Beltz.
- Jurecka, A. & Hartig, J. (2007). Computer- und netzwerkbasierendes Assessment. In J. Hartig & E. Klieme (Hrsg.), *Möglichkeiten und Voraussetzungen technologiebasierter Kompetenzdiagnostik. Eine Expertise im Auftrag des Bundesministeriums für Bildung und Forschung* (Bildungsforschung Bd. 20, S. 37-48). Berlin: BMBF.

- Kaminski, H. (2006). Wie viel Politik braucht die ökonomische Bildung? In G. Weißeno (Hrsg.), *Politik und Wirtschaft unterrichten* (S. 144-160). Wiesbaden: VS.
- Kemna, P. W. (2012). *Messung pädagogischer Basiskompetenzen von Lehrerinnen und Lehrern. Entwicklung von Testinstrumenten*. Münster: Waxmann.
- Klauer, K. J. & Leutner, D. (2012). *Lehren und Lernen. Einführung in die Instruktionspsychologie* (2. Aufl.). Weinheim: Beltz PVU.
- Kleickmann, T., Richter, D., Kunter, M., Elsner, J., Besser, M., Krauss, S. et al. (2013). Teachers' content knowledge and pedagogical content knowledge: The role of structural differences in teacher education. *Journal of Teacher Education*, 64(1), 90-106.
- Klein, S. P., Stecher, B. M., Shavelson, R. J., McCaffrey, D., Ormseth, T., Bell, R. M. et al. (1998). Analytic versus holistic scoring of science performance tasks. *Applied Measurement in Education*, 11(2), 121-137.
- Klieme, E., Avenarius, H., Blum, W., Döbrich, P., Gruber, H., Prenzel, M. et al. (2007, 2009). *Zur Entwicklung nationaler Bildungsstandards. Eine Expertise*. Bonn: Bundesministerium für Bildung und Forschung. Abgerufen am 18.04.2015 unter http://www.bmbf.de/pub/zur_entwicklung_nationaler_bildungsstandards.pdf
- Klieme, E. & Hartig, J. (2007). Kompetenzkonzepte in den Sozialwissenschaften und im empirischen Diskurs. In M. Prenzel, I. Gogolin & H.-H. Krüger (Hrsg.), *Kompetenzdiagnostik* (Zeitschrift für Erziehungswissenschaft, Sonderheft 8, S. 11-29). Wiesbaden: VS.
- Klieme, E. & Leutner, D. (2006). Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen. Beschreibung eines neu eingerichteten Schwerpunktprogramms der DFG. *Zeitschrift für Pädagogik*, 52(6), 876-903.
- Klieme, E., Lipowsky, F., Rakoczy, K. & Ratzka, N. (2006). Qualitätsdimensionen und Wirksamkeit von Mathematikunterricht. Theoretische Grundlagen und ausgewählte Ergebnisse des Projekts "Pythagoras". In M. Prenzel & L. Allolio-Näcke (Hrsg.), *Untersuchungen zur Bildungsqualität von Schule. Abschlussbericht des DFG-Schwerpunktprogramms* (S. 127-146). Münster: Waxmann.
- Klieme, E., Maag-Merki, K. & Hartig, J. (2007). Kompetenzbegriff und Bedeutung von Kompetenzen im Bildungswesen. In J. Hartig & E. Klieme (Hrsg.), *Möglichkeiten und Voraussetzungen technologiebasierter Kompetenzdiagnostik. Eine Expertise im*

- Auftrag des Bundesministeriums für Bildung und Forschung* (Bildungsforschung Bd. 20, S. 5-15). Berlin: BMBF.
- Klieme, E. & Rakoczy, K. (2008). Empirische Unterrichtsforschung und Fachdidaktik. Outcome-orientierte Messung und Prozessqualität des Unterrichts. *Zeitschrift für Pädagogik*, 54(2), 222-237.
- KMK (2004). *Standards für die Lehrerbildung: Bericht der Arbeitsgruppe*. Abgerufen am 05.03.2015 unter http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2004/2004_12_16-Standards_Lehrerbildung-Bericht_der_AG.pdf
- KMK (2014). *Standards für die Lehrerbildung: Bildungswissenschaften (Beschluss der Kultusministerkonferenz vom 16.12.2004 i. d. F. vom 12.06.2014)*. Abgerufen am 05.03.2015 unter http://www.akkreditierungsrat.de/fileadmin/Seiteninhalte/KMK/Vorgaben/KMK_Lehrerbildung_Standards_Bildungswissenschaften_aktuell.pdf
- KMK (2015). *Ländergemeinsame inhaltliche Anforderungen für die Fachwissenschaften und Fachdidaktiken in der Lehrerbildung (Beschluss der Kultusministerkonferenz vom 16.10.2008 i.d.F. vom 12.02.2015)*. Abgerufen am 16.04.2015 unter http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2008/2008_10_16-Fachprofile-Lehrerbildung.pdf
- Knievel, I., Lindmeier, A. M. & Heinze, A. (2015). Beyond knowledge: Measuring primary teachers' subject-specific competences in and for teaching mathematics with items based on video vignettes. *International Journal of Science and Mathematics Education*, 13, 309-329.
- Kohn, W. & Öztürk, R. (2013). *Statistik für Ökonomen. Datenanalyse mit R und SPSS* (2. Aufl.). Berlin: Springer.
- Kölbach, E. & Sumfleth, E. (2011). Kontexteinflüsse beim Lernen mit Beispielaufgaben im Fach Chemie. In S. Bernholt (Hrsg.), *Konzepte fachdidaktischer Strukturierung für den Unterricht* (Gesellschaft für Didaktik der Chemie und Physik, Bd. 32, S. 254-256). Berlin: LIT.
- Kölbach, E. & Sumfleth, E. Analyse von Kontexteffekten beim Lernen mit Lösungsbeispielen im Fach Chemie. *Zeitschrift für Didaktik der Naturwissenschaften*, 19, 159-188.
- König, J. (2010). Lehrerprofessionalität – Konzepte und Ergebnisse der internationalen und deutschen Forschung am Beispiel fachübergreifender, pädagogischer Kompetenzen. In J. König & B. Hofmann (Hrsg.), *Professionalität von Lehrkräften – Was sollen*

- Lehrkräfte im Lese- und Schreibunterricht wissen und können?* (S. 40-105). Berlin: DGLS. Abgerufen am 16.01.2015 unter https://www.hf.uni-koeln.de/data/eso/File/Koenig/Beitrag_Koenig_2010.pdf
- König, J., Blömeke, S. & Doll, J. (2011). Pädagogisches Wissen von Deutsch-, Englisch- und Mathematiklehramtsstudierenden. In S. Blömeke, A. Bremerich-Vos, H. Haudeck, G. Kaiser, G. Nold, K. Schwippert & H. Willenberg (Hrsg.), *Kompetenzen von Lehramtsstudierenden in gering strukturierten Domänen. Erste Ergebnisse aus TEDS-LT* (S. 135-157). Münster: Waxmann.
- König, J., Blömeke, S., Klein, P., Suhl, U., Busse, A. & Kaiser, G. (2014). Is teachers' general pedagogical knowledge a premise for noticing and interpreting classroom situations? A video-based assessment approach. *Teaching and Teacher Education*, 38, 76-88.
- Krauss, S., Baumert, J. & Blum, W. (2008a). Secondary mathematics teachers' pedagogical content knowledge and content knowledge: Validation of the COACTIV constructs. *Mathematics Education*, 40, 873-892.
- Krauss, S., Blum, W., Brunner, M., Neubrand, M., Baumert, J., Kunter, M. et al. (2011). Konzeptualisierung und Testkonstruktion zum fachbezogenen Professionswissen von Mathematiklehrkräften. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand, *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV* (S. 135-161). Münster: Waxmann.
- Krauss, S. & Bruckmaier, G. (2014). Das Experten-Paradigma in der Forschung zum Lehrerberuf. In E. Terhart, H. Bennewitz & M. Rothland (Hrsg.), *Handbuch der Forschung zum Lehrerberuf* (2. Aufl., S. 241-261). Münster: Waxmann.
- Krauss, S., Neubrand, M., Blum, W., Baumert, J., Brunner, M., Kunter, M. et al. (2008b). Die Untersuchung des professionellen Wissens deutscher Mathematik-Lehrerinnen und -Lehrer im Rahmen der COACTIV-Studie. *Journal für Mathematik-Didaktik*, 29(3/4), 223-258.
- Kunter, M. & Baumert, J. (2011). Das COACTIV-Forschungsprogramm zur Untersuchung professioneller Kompetenz von Lehrkräften – Zusammenfassung und Diskussion. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV* (S. 345-366). Münster: Waxmann.

- Kunter, M., Baumert, J., Blum, W., Klusmann, U., Krauss, S. & Neubrand, M. (Hrsg.) (2011). *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV*. Münster: Waxmann.
- Landesinitiative FÜR politische Bildung (2006). *Bildungspolitische Wende: Schulministerium gegen Politikunterricht an Gymnasien? Initiative FÜR politische Bildung gegründet*. Abgerufen am 12.03.2015 unter <http://www.uni-bielefeld.de/soz/ag/hedtke/pdf/pressemappeV0.7.PDF>
- Lange, D. (2011). Konzepte als Grundlage der politischen Bildung. Lerntheoretische und fachdidaktische Überlegungen. In Autorengruppe Fachdidaktik (Hrsg.), *Konzepte der politischen Bildung. Eine Streitschrift* (Reihe Politik und Bildung, Bd. 64, S. 95-109). Schwalbach/Ts.: Wochenschau.
- Langer, K., Schulz von Thun, F. & Tausch, R. (1974). *Verständlichkeit in Schule, Verwaltung, Politik und Wissenschaft* (2. Aufl.). München: Reinhardt.
- Langner, F. (2013). Theoriegeleitete Praxis als Leitbild beruflichen Handelns – Reflexionen zu dem eigenen Professionsverständnis. In K.-P. Hufer & D. Richter (Hrsg.), *Politische Bildung als Profession. Verständnisse und Forschungen* (Perspektiven politischer Bildung, Schriftenreihe Bd. 1355, S. 135-145). Bonn: bpb.
- Lehmann-Grube, S. K. & Nickolaus, R. (2009). Professionalität als kognitive Disposition. In O. Zlatkin-Troitschanskaia, K. Beck, D. Sembill, R. Nickolaus & R. Mulder (Hrsg.), *Lehrerprofessionalität. Bedingungen, Genese, Wirkungen und ihre Messung* (S. 59-70). Weinheim: Beltz.
- Leinhardt, G. (2010). Introduction: Explaining instructional explanations. In M. K. Stein & L. Kucan (Hrsg.), *Instructional explanations in the disciplines* (S. 1-5). New York: Springer.
- Leuders, T. & Holzäpfel, L. (2011). Kognitive Aktivierung im Mathematikunterricht. *Unterrichtswissenschaft*, 39(3), 213-230.
- Leuders, T. & Sodian, B. (2013). Inwiefern sind Kompetenzmodelle dazu geeignet kognitive Prozesse von Lernenden zu beschreiben? *Zeitschrift für Erziehungswissenschaft*, 16, 27-33.
- Lienert, G. A. & Raatz, U. (1998). *Testaufbau und Testanalyse* (6. Aufl.). Weinheim: Psychologie Verlags Union.
- Lindmeier, A. (2013). Video-vignettenbasierte standardisierte Erhebung von Lehrerkognitionen. In U. Riegel & K. Macha (Hrsg.), *Videobasierte*

- Kompetenzforschung in den Fachdidaktiken* (Fachdidaktische Forschungen, Bd. 4, S. 45-61). Münster: Waxmann.
- Lindmeier, A. (2011). *Modeling and measuring knowledge and competencies of teachers. A threefold domain-specific model for mathematics*. Münster: Waxmann.
- Lindmeier, A. (2014a). Validität von Maßen zur Erhebung von fachspezifischer Lehrerkognition (Moderierte Sektion). In J. Roth & J. Ames (Hrsg.), *Beiträge zum Mathematikunterricht 2014* (S. 67-68). Münster: WTM.
- Lindmeier, A. (2014b). *Video- und computerbasierte Kompetenz- und Instruktionsdiagnostik mit dem Softwarepaket vKid*. Abgerufen am 31.08.2015 unter <http://www.math.ipn.uni-kiel.de/vKID/flyer.pdf>
- Lindmeier, A. M., Heinze, A. & Reiss, K. (2012). Eine Machbarkeitsstudie zur Operationalisierung aktionsbezogener Kompetenz von Mathematiklehrkräften mit videobasierten Maßen. *Journal für Mathematik-Didaktik*, 34(1), 99-119.
- Lüdtke, O., Robitzsch, A., Trautwein, U. & Köller, O. (2007). Umgang mit fehlenden Werten in der psychologischen Forschung. *Psychologische Rundschau*, 58(2), 103-117.
- Maag Merki, K. & Werner, S. (2014). Erfassung und Bewertung professioneller Kompetenz von Lehrpersonen. In E. Terhart, H. Bennewitz & M. Rothland (Hrsg.), *Handbuch der Forschung zum Lehrerberuf* (2. Aufl., S. 745-763). Münster: Waxmann.
- Manzel, S. (2007). *Kompetenzzuwachs im Politikunterricht. Ergebnisse einer Interventionsstudie zum Kernkonzept Europa*. Münster: Waxmann.
- Manzel, S. (2013). Performanz im Politikunterricht – Theorieansätze und Entwicklungstrends zum professionellen Lehrer/-innenwissen und Unterrichtshandeln. In K.-P. Hufer & D. Richter (Hrsg.), *Politische Bildung als Profession. Verständnisse und Forschungen* (Perspektiven politischer Bildung, Schriftenreihe Bd. 1355, S. 203-216). Bonn: bpb.
- Manzel, S. (2014). Was wissen wir wirklich über politische Bildung in der Schule? Eine kritische Würdigung zum Stand der Disziplin. In D. Lange & T. Oeftering (Hrsg.), *Politische Bildung als lebenslanges Lernen* (Schriftenreihe der Gesellschaft für Politikdidaktik und politische Jugend- und Erwachsenenbildung, Bd. 13, S. 55-68). Schwalbach/Ts.: Wochenschau.
- Manzel, S. & Gronostay, D. (2013). Videografie im Politikunterricht. Erste Ergebnisse einer Pilotstudie zu domänenspezifischen Basisdimensionen. In U. Riegel & K. Macha

- (Hrsg.), *Videobasierte Kompetenzforschung in den Fachdidaktiken* (Fachdidaktische Forschungen, Bd. 4, S. 199-216). Münster: Waxmann.
- Manzel, S. & Neumann, D. (2015). Politikunterricht *anders* planen: Erkenntnisse der Fachdidaktik und empirischen Bildungsforschung zusammengedacht nutzen. In S. Frech & D. Richter (Hrsg.), *Politikunterricht professionell planen* (S. 69-83). Schwalbach/Ts.: Wochenschau.
- Massing, P. (2005). In Gesprächen lernen: Gesprächsformen in der politischen Bildung. In W. Sander (Hrsg.), *Handbuch politische Bildung* (Reihe Politik und Bildung, Bd. 32, 3. Aufl., S. 498-508). Schwalbach/Ts.: Wochenschau.
- Massing, P. (2013). "Jeder macht seins?" – Folgen der Bildungspolitik für die universitäre Lehrerbildung und das Lehramtsstudium von Politiklehrer/-innen. In K.-P. Hufer & D. Richter (Hrsg.), *Politische Bildung als Profession. Verständnisse und Forschungen* (Perspektiven politischer Bildung, Schriftenreihe Bd. 1355, S. 265-277). Bonn: bpb.
- Massing, P. & Weißeno, G. (1995). *Politik als Kern der politischen Bildung. Wege zur Überwindung unpolitischen Unterrichts*. Opladen: Leske + Budrich.
- May, M. (2007). *Demokratiefähigkeit und Bürgerkompetenzen. Kompetenztheoretische und normative Grundlagen der politischen Bildung* (Studien zur Schul- und Bildungsforschung, Bd. 26). Wiesbaden: VS.
- May, M. (2014). Politische Bildung als Beruf – Oder: Welche professionellen Herausforderungen stellen politische Bildungsprozesse an die Lehrenden? *Gesellschaft Wirtschaft Politik*, 4, 585-596.
- Mayer, J. & Wellnitz, N. (2014). Die Entwicklung von Kompetenzstrukturmodellen. In D. Krüger, I. Parchmann & H. Schecker (Hrsg.), *Methoden in der naturwissenschaftsdidaktischen Forschung* (S. 19-29). Berlin: Springer.
- Mayr, J. (2014). Der Persönlichkeitsansatz in der Forschung zum Lehrerberuf. Konzepte, Befunde und Folgerungen. In E. Terhart, H. Bennewitz & M. Rothland (Hrsg.), *Handbuch der Forschung zum Lehrerberuf* (2. Aufl., S. 189-215). Münster: Waxmann.
- Mayring, P. (2010). *Qualitative Inhaltsanalyse. Grundlagen und Techniken* (11. Aufl). Weinheim: Beltz.
- Meyer, H. (2011). *Was ist guter Unterricht?* (8. Aufl.). Berlin: Cornelsen.

- Moosbrugger, H. & Kelava, A. (2012). *Testtheorie und Fragebogenkonstruktion* (2. Aufl.). Berlin: Springer Medizin.
- MSW (Hrsg.) (2007). *Politik/Wirtschaft. Kernlehrplan für das Gymnasium – Sekundarstufe I (G8) in Nordrhein-Westfalen* (Heft 3429). Frechen: Ritterbach. Abgerufen am 18.04.2015 unter http://www.schulentwicklung.nrw.de/lehrplaene/upload/lehrplaene_download/gymnasium_g8/gym8_politik-wirtschaft.pdf
- MSW (Hrsg.) (2011). *Gesellschaftslehre. Erdkunde, Geschichte/Politik. Kernlehrplan für die Hauptschule in Nordrhein-Westfalen* (Heft 3202). Düsseldorf: Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen. Abgerufen am 18.04.2015 unter http://www.schulentwicklung.nrw.de/lehrplaene/upload/lehrplaene_download/hauptschule/GL_HS_KLP_Endfassung.pdf
- MSW (Hrsg.) (2014a). *Sozialwissenschaften und Sozialwissenschaften/Wirtschaft. Kernlehrplan für die Sekundarstufe II Gymnasium/Gesamtschule in Nordrhein-Westfalen* (Heft 4717). Düsseldorf: Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen. Abgerufen am 18.04.2015 unter http://www.schulentwicklung.nrw.de/lehrplaene/upload/klp_SII/sw/KLP_GOSt_SoWi.pdf
- MSW (2014b). *Praxiselemente in den lehramtsbezogenen Studiengängen. Runderlass des Ministeriums für Schule und Weiterbildung vom 28.06.2012 i.d.F. vom 15.06.2014*. Düsseldorf: Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen. Abgerufen am 18.04.2015 unter <http://www.schulministerium.nrw.de/docs/Recht/LAusbildung/Studium/Regelungen-Lehramtsstudium/Praxiselemente.pdf>
- MSW (o. J. a). *Gesellschaftslehre. Erdkunde, Geschichte, Politik. Kernlehrplan für die Gesamtschule – Sekundarstufe I in Nordrhein-Westfalen*. Düsseldorf: Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen. Abgerufen am 18.04.2015 unter http://www.schulentwicklung.nrw.de/lehrplaene/upload/lehrplaene_download/gesamtschule/GE_Gesellschaftslehre_Endfassung.pdf
- MSW (o. J. b). *Politik. Kernlehrplan für die Realschule in Nordrhein-Westfalen*. Düsseldorf: Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen. Abgerufen am 18.04.2015 unter http://www.schulentwicklung.nrw.de/lehrplaene/upload/klp_SI/RS/PL/RS_Politik_Endfassung.pdf

- MSWF (Hrsg.) (2001). *Rahmenvorgabe Politische Bildung*. Frechen: Ritterbach. Abgerufen am 18.04.2015 unter http://www.berufsbildung.schulministerium.nrw.de/cms/upload/_lehrplaene/a/uebergreifende_richtlinien/politische_bildung_500.pdf
- NBPTS (2002). *What teachers should know and be able to do. National Board for Professional Teaching Standards*. Abgerufen am 10.03.2015 unter http://www.nbpts.org/sites/default/files/what_teachers_should_know.pdf
- Neumann, D. (2014). „Ich erkenne Fehlkonzepte meiner Schüler/-innen und kann sie beim Lernen unterstützen“ – Diagnosekompetenz für den Politikunterricht modellieren. In S. Manzel (Hrsg.), *Politisch mündig werden. Politikkompetenz in der Schule aufbauen und diagnostizieren* (Schriften zur Didaktik der Sozialwissenschaften in Theorie und Unterrichtspraxis, Bd. 2, S. 49-60). Opladen: Barbara Budrich.
- Neumann, D. (2015). *Dokumentation der Studie „Politikdidaktisches Wissen und Können – eine explorative Studie zur Modellierung und Operationalisierung handlungsnaher Aspekte professioneller Kompetenzen“*. Organisation, Erhebungsinstrumente, Ratingmanuale und Datenaufbereitung. Unveröffentlichtes Manuskript, Universität Duisburg-Essen.
- Neumann, K. (2013). Mit welchem Auflösungsgrad können Kompetenzen modelliert werden? In welche Beziehung stehen Modelle zueinander, die Kompetenz in einer Domäne mit unterschiedlichem Auflösungsgrad beschreiben? *Zeitschrift für Erziehungswissenschaft*, 16, 35-39.
- Neuweg, G. H. (2002). Lehrerhandeln und Lehrerbildung im Lichte des Konzepts des impliziten Wissens. *Zeitschrift für Pädagogik*, 48(1), 10-29.
- Neuweg, G. H. (2014). Das Wissen der Wissensvermittler. Problemstellungen, Befunde und Perspektiven der Forschung zum Lehrerwissen. In E. Terhart, H. Bennewitz & M. Rothland (Hrsg.), *Handbuch der Forschung zum Lehrerberuf* (2. Aufl., S. 583-614). Münster: Waxmann.
- Newhouse, C. P., Lane, J. & Brown, C. (2007). Reflecting on teaching practices using digital video representation in teacher education. *Australian Journal of Teacher Education*, 32(3), 1-12.
- Newell, A. & McDermott, J. (1975). *PSG manual*. Pittsburgh: Department of Computer Science (Carnegie-Mellon University).
- Niggli, A. (2000). *Lernarrangements erfolgreich planen. Didaktische Anregungen zur Gestaltung offener Unterrichtsformen*. Aarau: Sauerländer.

- Oberle, M., Weschenfelder, E. & Weißeno, G. (2012). Professionskompetenz von Lehramtsstudierenden, Referendar/-innen und Lehrer/-innen. In I. Juchler (Hrsg.), *Unterrichtsleitbilder in der politischen Bildung* (Schriftenreihe der Gesellschaft für Politikdidaktik und politische Jugend- und Erwachsenenbildung, Bd. 11, S. 127-138). Schwalbach/Ts.: Wochenschau.
- Oberle, M., Weschenfelder, E. & Weißeno, G. (2013). Motivationale Orientierungen angehender und praktizierender Politiklehrkräfte. In A. Besand (Hrsg.), *Lehrer- und Schülerforschung in der politischen Bildung* (Schriftenreihe der Gesellschaft für Politikdidaktik und politische Jugend- und Erwachsenenbildung, Bd. 12, S. 55-67). Schwalbach/Ts.: Wochenschau.
- Oberle, M., Weschenfelder, E. & Weißeno, G. (2014). Beliefs als Element professioneller Kompetenz bei Politiklehrkräften in Deutschland. In B. Ziegler (Hrsg.), *Vorstellungen, Konzepte und Kompetenzen von Lehrpersonen der politischen Bildung. Beiträge zur Tagung Politische Bildung empirisch 2012* (Politische Bildung in der Schweiz, Bd. 3, S. 124-137). Zürich: Rüegger.
- Offenberg, E. & Walke, J. (2013). Die Reform der Praxisphasen in der Ersten Phase der Lehrerbildung. Eine qualitative Dokumentenanalyse. Bonn: Stifterverband für die deutsche Wissenschaft.
- Oleschko, S. (2013). Sprachbildung und Sprachförderung im Politikunterricht – Argumente für sprachsensiblen Unterricht. In A. Besand (Hrsg.), *Lehrer- und Schülerforschung in der politischen Bildung* (Schriftenreihe der Gesellschaft für Politikdidaktik und politische Jugend- und Erwachsenenbildung, Bd. 12, S. 193-207). Schwalbach/Ts.: Wochenschau.
- Oser, F. (2001). Standards: Kompetenzen von Lehrpersonen. In F. Oser & J. Oelkers (Hrsg.), *Die Wirksamkeit der Lehrerbildungssysteme. Von der Allrounderbildung zur Ausbildung professioneller Standards* (Nationales Forschungsprogramm 33. Wirksamkeit unserer Bildungssysteme, S. 215-342). Chur: Rüegger.
- Oser, F., Heinzer, S. & Salzmann, P. (2010). Die Messung der Qualität von professionellen Kompetenzprofilen von Lehrpersonen mit Hilfe der Einschätzung von Filmvignetten. Chancen und Grenzen des advokatorischen Ansatzes. *Unterrichtswissenschaft*, 38(1), 5-28.
- Oswald, M. E. & Gadenne, V. (1984). Wissen, Können und künstliche Intelligenz. *Sprache und Kognition*(3), S. 173-184.

- Pant, H. A. (2013). Wer hat einen Nutzen von Kompetenzmodellen? *Zeitschrift für Erziehungswissenschaft*, 16(1), 71-79.
- Paris, S. G., Lipson, M. Y. & Wixson, K. K. (1983). Becoming a strategic reader. *Contemporary Educational Psychology*, 8, 293-316.
- Park, S. & Oliver, J. S. (2008). Revisiting the conceptualisation of pedagogical content knowledge (PCK): PCK as a conceptual tool to understand teachers as professionals. *Research in Science Education*, 38, 261-284.
- Petko, D., Waldis, M., Pauli, C. & Reusser, K. (2003). Methodologische Überlegungen zur videogestützten Forschung in der Mathematikdidaktik. Ansätze der TIMSS 1999 Video Studie und ihrer schweizerischen Erweiterung. *Zentralblatt für Didaktik der Mathematik*, 35(6), 265-280.
- Petrik, A. (2009). „... aber das klappt nicht in der Schulpraxis!“ Skizze einer kompetenz- und fallorientierten Hochschuldidaktik für die Politiklehrer-Ausbildung. *Journal of Social Science Education*, 8(2), 57-80.
- Petrik, A. (2012). Der heimliche politikdidaktische Kanon. In I. Juchler (Hrsg.), *Unterrichtsleitbilder in der politischen Bildung* (Schriftenreihe der Gesellschaft für Politikdidaktik und politische Jugend- und Erwachsenenbildung, Bd. 11, S. 71-85). Schwalbach/Ts.: Wochenschau.
- Pfeiffer, T. (2008). Wissensstrukturen. In G. Weißeno (Hrsg.), *Politikkompetenz. Was Unterricht zu leisten hat* (Schriftenreihe Bd. 645, S. 76-88). Bonn: bpb.
- Praetorius, A.-K. (2014). *Messung von Unterrichtsqualität durch Ratings* (Pädagogische Psychologie und Entwicklungspsychologie, Bd. 90). Münster: Waxmann.
- Rehm, M. & Bölsterli, K. (2014). Entwicklung von Unterrichtsvignetten. In D. Krüger, I. Parchmann & H. Schecker (Hrsg.), *Methoden in der naturwissenschaftsdidaktischen Forschung* (S. 213-225). Berlin: Springer.
- Reinhardt, S. (2009). Gelingende Lehrerbildung – Professionstheorie und Fachdidaktik, Erfahrungen und Konsequenzen. *Journal of Social Science Education*, 8(2), S. 23-31.
- Reinhardt, S. (2011). Fachdidaktische Prinzipien als Brücken zwischen Gegenstand und Methode: Unterrichtsplanung. In Autorengruppe Fachdidaktik (Hrsg.), *Konzepte der politischen Bildung. Eine Streitschrift* (Reihe Politik und Bildung, Bd. 64, S. 147-162). Schwalbach/Ts.: Wochenschau.
- Reinhardt, S. (2012a). Fachdidaktische Prinzipien als Unterrichtsleitbilder in der politischen Bildung. In I. Juchler (Hrsg.), *Unterrichtsleitbilder in der politischen Bildung*

- (Schriftenreihe der Gesellschaft für Politikdidaktik und politische Jugend- und Erwachsenenbildung, Bd. 11, S. 35-45). Schwalbach/Ts.: Wochenschau.
- Reinhardt, S. (2012b). *Politikdidaktik. Praxishandbuch für die Sekundarstufe I und II* (4. Aufl.). Berlin: Cornelsen.
- Rentzsch, K. & Schütz, K. (2009). *Psychologische Diagnostik. Grundlagen und Perspektiven* (Grundriss der Psychologie, Bd. 16). Stuttgart: Kohlhammer.
- Reusser, K. (2009). Unterricht. In S. Andresen, R. Casale, T. Gabriel, R. Horlacher, S. Larcher Klee & J. Oelkers (Hrsg.), *Handwörterbuch Erziehungswissenschaft* (S. 881-896). Basel: Beltz.
- Reusser, K. & Pauli, C. (2014). Berufsbezogene Überzeugungen von Lehrerinnen und Lehrern. In E. Terhart, H. Bennewitz & M. Rothland (Hrsg.), *Handbuch der Forschung zum Lehrerberuf* (2. Aufl., S. 642-661). Münster: Waxmann.
- Riese, J. (2009). *Professionelles Wissen und professionelle Handlungskompetenz von (angehenden) Physiklehrkräften*. Berlin: Logos.
- Riese, J. & Reinhold, P. (2010). Empirische Erkenntnisse zur Struktur professioneller Handlungskompetenz von angehenden Physiklehrkräften. *Zeitschrift für Didaktik der Naturwissenschaften*, 16, 167-187.
- Riese, J. & Reinhold, P. (2012). Die professionelle Kompetenz angehender Physiklehrkräfte in verschiedenen Ausbildungsreformen. Empirische Hinweise für eine Verbesserung des Lehramtsstudiums. *Zeitschrift für Erziehungswissenschaft*, 15, 111-143.
- Riese, J. & Reinhold, P. (2014). Entwicklung eines Leistungstests für fachdidaktisches Wissen. In D. Krüger, I. Parchmann & H. Schecker (Hrsg.), *Methoden in der naturwissenschaftsdidaktischen Forschung* (S. 257-267). Berlin: Springer.
- Robitzsch, A. (2013). Wie robust sind Struktur- und Niveaumodelle? Wie zeitlich stabil und über Situationen hinweg konstant sind Kompetenzen? *Zeitschrift für Erziehungswissenschaft*, 16, 41-45.
- Ryle, G. (1949, 2000). *The concept of mind*. London: Penguin Books.
- Sander, W. (2005). Theorie der politischen Bildung: Geschichte – didaktische Konzeptionen – aktuelle Tendenzen und Probleme. In W. Sander (Hrsg.), *Handbuch politische Bildung* (Reihe Politik und Bildung, Bd. 32, 3. Aufl., S. 13-47). Schwalbach/Ts.: Wochenschau.

- Sander, W. (2008). *Politik entdecken – Freiheit leben. Didaktische Grundlagen politischer Bildung* (Reihe Politik und Bildung, Bd. 50, 3. Aufl.). Schwalbach/Ts.: Wochenschau.
- Sander, W. (2011). Kompetenzorientierung in Schule und politischer Bildung – eine kritische Zwischenbilanz. In Autorengruppe Fachdidaktik (Hrsg.), *Konzept der politischen Bildung. Eine Streitschrift* (Reihe Politik und Bildung, Bd. 64, S. 9-25). Schwalbach/Ts.: Wochenschau.
- Sander, W. (2013). Die Kompetenzblase – Transformationen und Grenzen der Kompetenzorientierung. *Zeitschrift für Didaktik der Gesellschaftswissenschaften*, 1, 100-124.
- Schelle, C. (2005). Adressatenorientierung. In W. Sander (Hrsg.), *Handbuch politische Bildung* (Reihe Politik und Bildung, Bd. 32, 3. Aufl., S. 79-92). Schwalbach/Ts.: Wochenschau.
- Scherb, A. (2002). *Ist eine konstruktivistische Politikdidaktik möglich?* (Berichte aus der Pädagogik). Aachen: Shaker.
- Schmelzing, S. (2010). *Das fachdidaktische Wissen von Biologielehrkräften: Konzeptionalisierung, Diagnostik, Struktur und Entwicklung im Rahmen der Biologielehrerbildung*. Berlin: Logos.
- Schmelzing, S., Wüsten, S., Sandmann, A. & Neuhaus, B. (2010). Fachdidaktisches Wissen und Reflektieren im Querschnitt der Biologielehrerbildung. *Zeitschrift für Didaktik der Naturwissenschaften*, 16, 189-207.
- Schmidt-Atzert, L. & Amelang, M. (2012). *Psychologische Diagnostik* (5 Aufl.). Berlin: Springer.
- Schmiederer, R. (1977). *Politische Bildung im Interesse der Schüler*. Köln: Europäische Verlagsanstalt.
- Schmitt, N. (1996). Uses and abuses of coefficient alpha. *Psychological Assessment*, 8(4), 350-353.
- Sedlmeier, P. & Renkewitz, F. (2008). *Forschungsmethoden und Statistik in der Psychologie*. München: Pearson.
- Seifried, J. & Ziegler, B. (2009). Domänenbezogene Professionalität. In O. Zlatkin-Troitschanskaia, K. Beck, D. Sembill, R. Nickolaus & R. Mulder (Hrsg.), *Lehrerprofessionalität. Bedingungen, Genese, Wirkungen und ihre Messung* (S. 83-92). Weinheim: Beltz.

- Sherin, M. G. & van Es, E. A. (2009). Effects of video club participation on teachers' professional vision. *Journal of Teacher Education*, 60(1), 20-37. Abgerufen am 27.09.2015 unter <http://jte.sagepub.com/content/60/1/20.full.pdf>
- Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15(2), 4-14. Abgerufen am 17.04.2015 unter <http://links.jstor.org/sici?sici=0013-189X%28198602%2915%3A2%3C4%3ATWUKGI%3E2.0.CO%3B2-X>
- Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, 57(1), 1-22.
- Smith, D. C. & Neale, D. C. (1989). The construction of subject matter knowledge in primary science teaching. *Teaching & Teacher Education*, 5(1), 1-20.
- Stender, A. (2014). *Unterrichtsplanung. Vom Wissen zum Handeln. Theoretische Entwicklung und empirische Überprüfung des Transformationsmodells der Unterrichtsplanung*. Dissertation, Christian-Albrechts-Universität zu Kiel.
- Stern, E. (2009). Implizite und explizite Lernprozesse bei Lehrerinnen und Lehrern. In O. Zlatkin-Troitschanskaia, K. Beck, D. Sembill, R. Nickolaus & R. Mulder (Hrsg.), *Lehrerprofessionalität. Bedingungen, Genese, Wirkungen und ihre Messung* (S. 355-364). Weinheim: Beltz.
- Szukala, A. (2013). Der Zusammenhang zwischen epistemologischen Überzeugungen und Lehr-Lernüberzeugungen in der sozialwissenschaftlichen Domäne. In A. Besand (Hrsg.), *Lehrer- und Schülerforschung in der politischen Bildung* (Schriftenreihe der Gesellschaft für Politikdidaktik und politische Jugend- und Erwachsenenbildung, Bd. 12, S. 33-54). Schwalbach/Ts.: Wochenschau.
- TUM (2015). *Videobasierte Kompetenz- und Instruktionsdiagnostik mit vKid (TUM-Wiki)*. Abgerufen am 31.08.2015 unter <https://vkid.wiki.tum.de/>
- Teddlie, C., Stringfield, S. & Burdett, J. (2003). International comparisons of the relationship among educational effectiveness, evaluation and improvement variables: An overview. *Journal of Personnel Evaluation in Education*, 17(1), 5-20.
- Tenorth, H.-E. (2006). Professionalität im Lehrerberuf. Ratlosigkeit der Theorie, gelingende Praxis. *Zeitschrift für Erziehungswissenschaft*, 9(4), 580-597.
- Tepner, O., Borowski, A., Dollny, S., Fischer, H. E., Jüttner, M., Kirschner, S. et al. (2012). Modell zur Entwicklung von Testitems zur Erfassung des Professionswissens von Lehrkräften in den Naturwissenschaften. *Zeitschrift für Didaktik der Naturwissenschaften*, 18, 7-28.

- Tergan, S.-O. (1986). *Modelle der Wissensrepräsentation als Grundlage qualitativer Wissensdiagnostik* (Beiträge zur psychologischen Forschung, Bd. 7). Wiesbaden: Springer Fachmedien.
- UDE (2014a). *Fachprüfungsordnung für das Fach Sozialwissenschaften im Bachelor-Studiengang mit Lehramtsoption Gymnasium/Gesamtschule an der Universität Duisburg-Essen. Vom 29. Juni 2012, zuletzt geändert durch dritte Änderungsordnung vom 24. März 2014*. Essen: Universität Duisburg-Essen. Abgerufen am 18.04.2014 unter https://www.uni-due.de/imperia/md/content/zentralverwaltung/bereinigte_sammlung/8_35_26.pdf
- UDE (2014b). *Fachprüfungsordnung für das Fach Sozialwissenschaften im Bachelor-Studiengang mit Lehramtsoption Haupt-, Real- und Gesamtschulen an der Universität Duisburg-Essen. Vom 29. Juni 2012, zuletzt geändert durch zweite Änderungsordnung vom 24. März 2014*. Essen: Universität Duisburg-Essen. Abgerufen am 18.04.2015 unter https://www.uni-due.de/imperia/md/content/zentralverwaltung/bereinigte_sammlung/8_35_24.pdf
- Van de Sande, C. & Greeno, J. G. (2010). A framing of instructional explanations: Let us explain *with* you. In M. K. Stein & L. Kucan (Hrsg.), *Instructional explanations in the disciplines* (S. 69-82). New York: Springer.
- Van Dijk, E. M. & Kattmann, U. (2007). A research model for the study of science teachers' PCK and improving teacher education. *Teaching and Teacher Education*, 23(6), 885-897.
- Van Dijk, E. M. & Kattmann, U. (2010). Evolution im Unterricht: Eine Studie über fachdidaktisches Wissen von Lehrerinnen und Lehrern. *Zeitschrift für Didaktik der Naturwissenschaften*, 16, 7-21.
- Van Driel, J. H., Verloop, N. & de Vos, W. (1998). Developing teacher' pedagogical content knowledge. *Journal of research in science teaching*, 35(6), 673-695.
- Vogelsang, C. & Reinhold, P. (2013). Zur Handlungsvalidität von Tests zum professionellen Wissen von Lehrkräften. *Zeitschrift für Didaktik der Naturwissenschaften*, 19, 103-128.
- Voss, T. & Kunter, M. (2011). Pädagogisch-psychologisches Wissen von Lehrkräften. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV* (S. 193-214). Münster: Waxmann.

- Wehling, H.-G. (1977). Konsens à la Beutelsbach? Nachlese zu einem Expertengespräch. In S. Schiele & H. Schneider (Hrsg.), *Das Konsensproblem in der politischen Bildung*. Stuttgart: Klett.
- Weinert, F. E. (2001). Vergleichende Leistungsmessung in Schulen – eine umstrittene Selbstverständlichkeit. In F. E. Weinert (Hrsg.), *Leistungsmessungen in Schulen* (2. Aufl., S. 17-31). Weinheim: Beltz.
- Weinert, F. E., Schrader, F.-W. & Helmke, A. (1990). Unterrichtsexpertise – Ein Konzept zur Verringerung der Kluft zwischen zwei theoretischen Paradigmen. In L.-M. Alisch, J. Baumert & K. Beck (Hrsg.), *Professionswissen und Professionalisierung* (Braunschweiger Studien zur Erziehungs- und Sozialarbeitswissenschaft, Bd. 28, S. 173-206). Braunschweig: TU Braunschweig.
- Weißeno, G. (2008). Politikkompetenz. Neue Aufgaben für Theorie und Praxis. In G. Weißeno (Hrsg.), *Politikkompetenz. Was Unterricht zu leisten hat* (Schriftenreihe Bd. 645, S. 11-20). Bonn: bpb.
- Weißeno, G. (2012). Zum Stand empirischer politikdidaktischer Forschung. In I. Juchler (Hrsg.), *Unterrichtsleitbilder in der politischen Bildung* (Schriftenreihe der Gesellschaft für Politikdidaktik und politische Jugend- und Erwachsenenbildung, Bd. 11, S. 115-126). Schwalbach/Ts.: Wochenschau.
- Weißeno, G., Detjen, J., Juchler, I., Massing, P. & Richter, D. (2010). *Konzepte der Politik – ein Kompetenzmodell* (Schriftenreihe Bd. 1016). Bonn: bpb.
- Weißeno, G. & Landwehr, B. (2015). Effektiver Unterricht über die Europäische Union – Ergebnisse einer Studie zur Schülerperzeption von Politikunterricht. In M. Oberle (Hrsg.), *Die Europäische Union erfolgreich vermitteln. Perspektiven der politischen EU-Bildung heute* (S. 99-109). Wiesbaden: Springer VS.
- Weißeno, G., Weschenfelder, E. & Oberle, M. (2013). Empirische Ergebnisse zur professionellen Kompetenz von Politiklehrer/-innen. In K.-P. Hufer & D. Richter (Hrsg.), *Politische Bildung als Profession. Verständnisse und Forschungen* (Perspektiven politischer Bildung, Schriftenreihe Bd. 1355, S. 187-202). Bonn: bpb.
- Weschenfelder, E. (2014a). *Professionelle Kompetenz von Politiklehrkräften. Eine Studie zu Wissen und Überzeugungen*. Wiesbaden: Springer VS.
- Weschenfelder, E. (2014b). Professionelles Wissen und Überzeugungen bei Politiklehrkräften. In D. Lange & T. Oeftering (Hrsg.), *Politische Bildung als lebenslanges Lernen* (Schriftenreihe der Gesellschaft für Politikdidaktik und

politische Jugend- und Erwachsenenbildung, Bd. 13, S. 149-159). Schwalbach/Ts.: Wochenschau.

- Weschenfelder, E., Weißeno, G. & Oberle, M. (2014). Professionelles Wissen angehender deutscher Politiklehrkräfte. In B. Ziegler (Hrsg.), *Vorstellungen, Konzepte und Kompetenzen von Lehrpersonen der politischen Bildung. Beiträge zur Tagung Politische Bildung empirisch 2012* (Politische Bildung in der Schweiz, Bd. 3, S. 138-155). Zürich: Rüegger.
- Wiechmann, J. (2011). Direkte Instruktion. In J. Wiechmann (Hrsg.), *Zwölf Unterrichtsmethoden. Vielfalt für die Praxis* (5. Aufl., S. 39-51). Weinheim: Beltz.
- Wilhelm, O. & Nickolaus, R. (2013). Was grenzt das Kompetenzkonzept von etablierten Kategorien wie Fähigkeit, Fertigkeit oder Intelligenz ab? *Zeitschrift für Erziehungswissenschaft*, 16, 23-26.
- Yore, L. D., Pimm, D. & Tuan, H.-L. (2007). The literacy component of mathematical and scientific literacy. *International Journal of Science and Mathematics Education*, 5(4), 559-589.
- Zimbardo, P. G. & Gerrig, R. J. (2004). *Psychologie* (16. Aufl.). München: Pearson.
- Zurstrassen, B. (2009). Kompetenzorientierte Lehrerbildung in den sozialwissenschaftlichen Unterrichtsfächern: Blühende Landschaften in der sozialwissenschaftlichen Lehrerbildung von morgen? *Journal of Social Science Education*, 8(2), 32-45.

10 Tabellen

Tabelle 1: Übersicht über Facetten fachdidaktischen Wissens in unterschiedlichen Studien (eigene Darstellung nach Schmelzing, 2010, S. 23).....	41
Tabelle 2: Operationalisierungen des fachdidaktischen Wissens über Instruktions- und Vermittlungsstrategien sowie über Kognitionen der Schüler/-innen in der mathematisch-naturwissenschaftlichen Domäne in unterschiedlichen Studien	43
Tabelle 3: Interne Konsistenz der (revidierten) Skalen sowie des Gesamttests (Reliabilitätskoeffizient α , Standardmessfehler SE und auf 16 Items geschätzter Alpha-Wert)	172
Tabelle 4: Durchschnittliche Testleistungen in den Skalen sowie im Gesamttest (kumulierte Testleistungen).....	175
Tabelle 5: Durchschnittliche Testleistungen in den Skalen sowie im Gesamttest (Personenfähigkeit)	175
Tabelle 6: Korrelationen zwischen den Skalen (Spearman's Rho).....	180
Tabelle 7: Übersicht über die durchschnittlichen Testleistungen (Personenfähigkeit) in den Skalen sowie im Gesamttest gruppiert nach Personenmerkmalen	181
Tabelle 8: Übersicht über die durchschnittlichen Testleistungen (Personenfähigkeit) in den revidierten Skalen sowie im revidierten Gesamttest gruppiert nach Personenmerkmalen ..	182
Tabelle 9: Ergebnisse des U-Tests auf Unterschiede in der zentralen Tendenz zwischen (Extrem-)Gruppen in den Skalen und im Gesamttest	183
Tabelle 10: Ergebnisse des U-Tests auf Unterschiede in der zentralen Tendenz zwischen (Extrem-)Gruppen in den revidierten Skalen und im revidierten Gesamttest.....	183

Tabelle A. 1: Beschreibung der Stichprobe: Abiturnote.....	236
Tabelle A. 2: Beschreibung der Stichprobe: Geschlecht und Alter	237
Tabelle A. 3: Beschreibung der Stichprobe: Studienabschluss, Schulform und Studienverlauf	238
Tabelle A. 4: Beschreibung der Stichprobe: Studierte politikdidaktische Modulteile der Studierenden gem. LABG 2009 (Bachelor of Arts)	239
Tabelle A. 5: Beschreibung der Stichprobe: Studierte politikdidaktische Modulteile der Studierenden gem. LABG 2002 (Erste Staatsprüfung)	239
Tabelle A. 6: Beschreibung der Stichprobe: Studierte politikdidaktische Modulteile beider Studierendengruppen (LABG 2002 und LABG 2009)	240
Tabelle A. 7: Beschreibung der Stichprobe: Absolvierte Praktika und Lehr-/Unterrichtserfahrungen der Studierenden gem. LABG 2009 (Bachelor of Arts).....	241
Tabelle A. 8: Beschreibung der Stichprobe: Ort und Fachbezug der Praktika der Studierenden gem. LABG 2009 (Bachelor of Arts)	242
Tabelle A. 9: Beschreibung der Stichprobe: Absolvierte Praktika und Lehr-/Unterrichtserfahrungen der Studierenden gem. LABG 2002 (Erste Staatsprüfung)..	242
Tabelle A. 10: Übersicht über technisch bedingt fehlende Werte im Computerprogramm vKid in der Skala zum deklarativen politikdidaktischen Wissen (FdW)	244
Tabelle A. 11: Übersicht über technisch bedingt fehlende Werte im Computerprogramm vKid in der Skala zum prozeduralen politikdidaktischen Wissen (FpW)	245
Tabelle A. 12: Übereinstimmungsmaße zwischen den Ratern in der Skala zum prozeduralen politikdidaktischen Wissen (FpW)	246
Tabelle A. 13: Itemkennwerte zur Skala zum Fachwissen (FW) mit Anzahl gültiger Fälle, Schwierigkeitsindex p_i , Itemtrennschärfe r_{it} sowie Verteilung der Testleistung (absolute Häufigkeiten, Mittelwert M und Standardabweichung SD)	247

Tabelle A. 14: Itemkennwerte zur Skala zum deklarativen politikdidaktischen Wissen (FdW) mit Anzahl gültiger Fälle, Schwierigkeitsindex p_i . Itemtrennschärfe r_{it} sowie Verteilung der Testleistung (absolute Häufigkeiten, Mittelwert M und Standardabweichung SD). 249

Tabelle A. 15: Itemkennwerte zur Skala zum prozeduralen politikdidaktischen Wissen mit Anzahl gültiger Fälle, Schwierigkeitsindex p_i . Itemtrennschärfe r_{it} sowie Verteilung der Testleistung (absolute Häufigkeiten, Mittelwert M und Standardabweichung SD). 250

11 Abbildungen

Abbildung 1: Angebots-Nutzungs-Modell der Wirkungsweise des Unterrichts (eigene Darstellung in Anlehnung an Helmke, 2014, S. 71).....	20
Abbildung 2: Arten des Gedächtnisses (eigene Darstellung in Anlehnung an Squire, 1987 zitiert nach Anderson, 2013, S. 160)	28
Abbildung 3: Strukturkomponenten und deren Zusammenhang in der ACT-Theorie (eigene Darstellung nach Anderson, 1983, S. 19).....	29
Abbildung 4: Konzepte des Lehrerwissens (eigene Darstellung in Anlehnung an Neuweg, 2014, S. 585)	32
Abbildung 5: Beispiel-Item zur Erfassung des lernprozessbezogenen politikdidaktischen Wissens in der PKP-Studie (eigene Darstellung nach Weschenfelder, 2014a, S. 171).....	37
Abbildung 6: Modell des Professionswissens von Lehrkräften in den Naturwissenschaften im Projekt ProwiN (eigene Darstellung in Anlehnung an Tepner et al., 2012, S. 19).....	53
Abbildung 7: Heuristisches Strukturmodell zum biologiedidaktischen Wissen (eigene Darstellung nach Schmelzing, 2010, S. 29)	54
Abbildung 8: Modell physikalisch-fachdidaktischen Wissens bei Riese (2009) (eigene Darstellung nach Gramzow et al., 2013, S. 13).....	61
Abbildung 9: Zusammenhang von Fähigkeiten und Kompetenzen (eigene Darstellung nach Eraut, 1998, S. 135).....	65
Abbildung 10: Darstellung der Entwicklungs- und Forschungsschritte von der Kompetenzmodellierung bis zur Kompetenzdiagnostik (eigene Darstellung in Anlehnung an Fleischer et al., 2013, S. 7).....	70
Abbildung 11: Grunddimensionen der Unterrichtsqualität und deren vermutliche Wirkung (eigene Darstellung in Anlehnung an Klieme, Lipowsky, Rakoczy & Ratzka, 2006, S. 131; Klieme & Rakoczy, 2008, S. 228).....	83

Abbildung 12: Allgemeines Kompetenzstrukturmodell professioneller Handlungskompetenzen (eigene Darstellung in Anlehnung an Baumert & Kunter, 2006, S. 482; Baumert & Kunter, 2011a, S. 32).....	102
Abbildung 13: Modell professioneller Handlungskompetenz von Politiklehrkräften der PKP-Studie (eigene Darstellung nach Weißen et al., 2013, S. 189).....	105
Abbildung 14: Mehrdimensionales Modell des Professionswissens von Politiklehrkräften (in Anlehnung an Tepner et al., 2012, S. 19 sowie ergänzend Baumert & Kunter, 2011a; Krauss et al., 2008b; Schmelzing, 2010; Weißen et al., 2013).....	112
Abbildung 15: Heuristisches mehrdimensionales Strukturmodell des politikdidaktischen Wissens (in Anlehnung an Schmelzing, 2010; Tepner et al., 2012; Weißen et al., 2013 sowie ergänzend Faissner et al., o. J., S. 7)	115
Abbildung 16: Aufbau der Testumgebung mit Laptop, akustischem Ausgabe-/Aufnahmegerät (Headset) und Notizklembrett mit Videoperipherie.....	131
Abbildung 17: Heuristisches mehrdimensionales Strukturmodell zum politikdidaktischen Wissen und Können mit Spezifizierung der Inhaltsdimension in der vorliegenden Studie (eigene Darstellung in Anlehnung an Schmelzing, 2010; Tepner et al., 2012, Weißen et al., 2013 sowie ergänzend Faissner et al., o. J., S. 14)	133
Abbildung 18: Beispiel-Item (geschlossenes Antwortformat) zur Erhebung des Fachwissens	136
Abbildung 19: Beispiel-Item (geschlossenes Antwortformat) zur Erhebung des deklarativen politikdidaktischen Wissens über Instruktions- und Vermittlungsstrategien im Computerprogramm vKid	138
Abbildung 20: Beispiel-Item (halb offenes Antwortformat) zur Erhebung des deklarativen politikdidaktischen Wissens über Instruktions- und Vermittlungsstrategien im Computerprogramm vKid	139

Abbildung 21: Beispiel-Item (geschlossenes Antwortformat) zur Erhebung des deklarativen politikdidaktischen Wissens über (Fehl-)Vorstellungen der Schüler/-innen im Computerprogramm vKid	140
Abbildung 22: Beispiel-Item zur Erhebung des prozeduralen politikdidaktischen Wissens (Aufgabentyp <i>unterstützender Impuls</i>) im Computerprogramm vKid.....	142
Abbildung 23: Beispieltranskript einer Rückmeldung eines Probanden (Prob_ID 3) in der Skala zum prozeduralen politikdidaktischen Wissen (Antworttyp <i>Erklärung</i>).....	150
Abbildung 24: Zusammenhang zwischen Itemschwierigkeit und Trennschärfe in den vorliegenden Daten zum Fachwissen (FW)	168
Abbildung 25: Zusammenhang zwischen Itemschwierigkeit und Trennschärfe in den vorliegenden Daten zum Fachwissen unter Berücksichtigung ausgeschlossener Items (FW_rev)	170
Abbildung 26: Streudiagramm zum Zusammenhang zwischen der kumulierten Bearbeitungszeit (in Millisekunden) und den durchschnittlichen Testleistungen der Probanden in der Skala zum prozeduralen politikdidaktischen Wissen (FpW)	177
Abbildung 27: Streudiagramm zum Zusammenhang zwischen der kumulierten Bearbeitungszeit (in Millisekunden) und den durchschnittlichen Testleistungen der Probanden in der revidierten Skala zum prozeduralen politikdidaktischen Wissen (FpW_rev)	178
Abbildung A. 1: Beschreibung der Stichprobe: Streudiagramm (gruppiert) zur Verteilung der Angaben zu den Lehr-/Unterrichtserfahrungen der Probanden in Zeitstunden.....	243
Abbildung A. 2: Zusammenhang zwischen Itemschwierigkeit und Trennschärfe in den vorliegenden Daten zur Skala zum deklarativen politikdidaktischen Wissen unter Berücksichtigung ausgeschlossener Items (FdW_rev).....	251

Abbildung A. 3: Zusammenhang zwischen Itemschwierigkeit und Trennschärfe in den vorliegenden Daten zur Skala zum prozeduralen politikdidaktischen Wissen unter Berücksichtigung ausgeschlossener Items (FpW_rev).....	252
Abbildung A. 4: Verteilung der durchschnittlichen Testleistungen der Probanden (Personenfähigkeit) in der Skala zum Fachwissen (FW) mit Normalverteilungskurve.....	253
Abbildung A. 5: Verteilung der durchschnittlichen Testleistungen der Probanden (Personenfähigkeit) in der revidierten Skala zum Fachwissen (FW_rev) mit Normalverteilungskurve.....	254
Abbildung A. 6: Verteilung der durchschnittlichen Testleistungen der Probanden (Personenfähigkeit) in der Skala zum deklarativen politikdidaktischen Wissen (FdW) mit Normalverteilungskurve.....	255
Abbildung A. 7: Verteilung der durchschnittlichen Testleistungen der Probanden (Personenfähigkeit) in der revidierten Skala zum deklarativen politikdidaktischen Wissen (FdW_rev) mit Normalverteilungskurve.....	256
Abbildung A. 8: Verteilung der durchschnittlichen Testleistungen der Probanden (Personenfähigkeit) in der Skala zum prozeduralen politikdidaktischen Wissen (FpW) mit Normalverteilungskurve.....	257
Abbildung A. 9: Verteilung der durchschnittlichen Testleistungen der Probanden (Personenfähigkeit) in der revidierten Skala zum prozeduralen politikdidaktischen Wissen (FpW_rev) mit Normalverteilungskurve.....	258
Abbildung A. 10: Histogramm mit Normalverteilungskurve zu den kumulierten Bearbeitungszeiten (in Millisekunden) der Probanden in der Skala zum prozeduralen politikdidaktischen Wissen (FpW)	259
Abbildung A. 11: Histogramm mit Normalverteilungskurve zu den kumulierten Bearbeitungszeiten (in Millisekunden) der Probanden in der revidierten Skala zum prozeduralen politikdidaktischen Wissen (FpW_rev)	260

Anhang

Tabelle A. 1: Beschreibung der Stichprobe: Abiturnote

Variable	Beschreibung	Wert	<i>N</i>	Prozent
Stichprobenumfang			50	
Abiturnote	Schulnote	1.0	1	2 %
		1.7	2	4 %
		1.8	3	6 %
		1.9	5	10 %
		2.0	3	6 %
		2.1	3	6 %
		2.2	7	14 %
		2.3	5	10 %
		2.4	2	4 %
		2.5	2	4 %
		2.6	3	6 %
		2.7	5	10 %
		2.9	2	4 %
		3.0	1	2 %
		3.1	2	4 %
		3.2	2	4 %
		3.3	1	2 %
		fehlend	1	2 %
Abiturnote	<i>M</i>	2.3		
	<i>SD</i>	0.47		

Anmerkung: Eine Probandin absolvierte ihren höchsten schulischen Bildungsabschluss nicht in Deutschland und machte daher keine Angaben zur Abiturnote (fehlender Wert).

Tabelle A. 2: Beschreibung der Stichprobe: Geschlecht und Alter

Variable	Beschreibung	Wert	<i>N</i>	Prozent
Stichprobenumfang			50	
Geschlecht	weiblich		29	58 %
	männlich		21	42 %
Alter	Jahre	30	2	4 %
		29	1	2 %
		28	1	2 %
		27	2	4 %
		26	4	8 %
		25	7	14 %
		24	8	16 %
		23	7	14 %
		22	15	30 %
		21	3	6 %
Alter	<i>M</i>	24		
	<i>SD</i>	2.25		

Tabelle A. 3: Beschreibung der Stichprobe: Studienabschluss, Schulform und Studienverlauf

Variable	Beschreibung	Wert	<i>N</i>	Prozent
Stichprobenumfang			50	
Studienabschluss	Bachelor of Arts (LABG 2009)		26	52 %
	Erste Staatsprüfung (LABG 2002)		24	48 %
Schulform	Haupt-, Real- und Gesamtschulen		12	24 %
	Gymnasien und Gesamtschulen		38	76 %
Studienverlauf	Fachsemester	1	1	2 %
		2	1	2 %
		4	1	2 %
		6	13	26 %
		7	11	22 %
		8	13	26 %
		9	1	2 %
		10	5	10 %
		12	2	4 %
		14	1	2 %
		20	1	2 %
Studienverlauf	<i>M</i>	8		
	<i>SD</i>	2.83		

Tabelle A. 4: Beschreibung der Stichprobe: Studierte politikdidaktische Moduleile der Studierenden gem. LABG 2009 (Bachelor of Arts)

Variable	Beschreibung	Wert	N	Prozent
Stichprobenumfang			26	
Modulteil- bzw. Veranstaltungsname	Grundlagen der Fachdidaktik der Sozialwissenschaften (LV_1)		4	15 %
	Lehren und Lernen in der sozialwissenschaftlichen Bildung (LV_2)		1	4 %
	Alle Veranstaltungen (LV_1 und LV_2)		19	73 %
		fehlend	2	8 %

Tabelle A. 5: Beschreibung der Stichprobe: Studierte politikdidaktische Moduleile der Studierenden gem. LABG 2002 (Erste Staatsprüfung)

Variable	Beschreibung	Wert	N	Prozent
Stichprobenumfang			24	
Modulteil- bzw. Veranstaltungsname	Grundlagen der Fachdidaktik der Sozialwissenschaften (LV_1)		3	13 %
	Lehren und Lernen in der sozialwissenschaftlichen Bildung (LV_2)		0	0 %
	Theorien und Modelle sozialwissenschaftlichen Unterrichts (LV_3)		0	0 %
	LV_1 und LV_2		0	0 %
	LV_1 und LV_3		2	8 %
	LV_2 und LV_3		1	4 %
	Alle Veranstaltungen (LV_1, LV_2 und LV_3)		18	75 %

Anmerkung: Die Modulteil- bzw. Veranstaltungsamen wurden zum Zwecke der besseren Vergleichbarkeit zwischen den Studierendengruppen nach Studienabschluss aufgrund der polyvalenten Studienstruktur und der vergleichbaren Inhalte sowie Kompetenzziele (vgl. Kapitel 7.1) an die entsprechenden Namen aus den Studiengängen gemäß LABG 2009 angepasst.

Tabelle A. 6: Beschreibung der Stichprobe: Studierte politikdidaktische Moduleile beider Studierendengruppen (LABG 2002 und LABG 2009)

Variable	Beschreibung	Wert	N	Prozent
Stichprobenumfang			50	
Modulteil- bzw. Veranstaltungsname	Grundlagen der Fachdidaktik der Sozialwissenschaften (LV_1)		7	14 %
	Lehren und Lernen in der sozialwissenschaftlichen Bildung (LV_2)		1	2 %
	Theorien und Modelle sozialwissenschaftlichen Unterrichts (LV_3)		0	0 %
	LV_1 und LV_2		19	38 %
	LV_1 und LV_3		2	4 %
	LV_2 und LV_3		1	2 %
	Alle Veranstaltungen (LV_1, LV_2 und LV_3)		18	36 %
		fehlend	2	4 %

Anmerkung: Die Modulteil- bzw. Veranstaltungsnamen wurden zum Zwecke der besseren Vergleichbarkeit zwischen den Studierendengruppen nach Studienabschluss aufgrund der polyvalenten Studienstruktur und der vergleichbaren Inhalte sowie Kompetenzziele (vgl. Kapitel 7.1) an die entsprechenden Namen aus den Studiengängen gemäß LABG 2009 angepasst. Die Lehrveranstaltung *Theorien und Modelle sozialwissenschaftlichen Unterrichts* ist in der Studienstruktur gemäß LABG 2009 im Studiengang Master of Education verortet und konnte aufgrund des sukzessiven Aufbaus der konsekutiven Bachelor-Master-Studienstruktur an der Universität Duisburg-Essen zum Befragungszeitpunkt von den Studierenden im Bachelor-Studiengang noch nicht besucht worden sein. Die Angaben zu dieser Veranstaltung beziehen sich daher ausschließlich auf die Studierenden gemäß LABG 2002.

Tabelle A. 7: Beschreibung der Stichprobe: Absolvierte Praktika und Lehr-/Unterrichtserfahrungen der Studierenden gem. LABG 2009 (Bachelor of Arts)

Variable	Beschreibung	Wert	<i>N</i>	Prozent
Stichprobenumfang			26	
Praktika	Eignungspraktikum (P_1)		1	4 %
	Orientierungspraktikum (P_2)		0	0 %
	Berufsfeldpraktikum (P_3)		0	0 %
	P_1 und P_2		1	4 %
	P_1 und P_3		0	0 %
	P_2 und P_3		4	15 %
	Alle Praktika (P_1, P_2 und P_3)		20	77 %
Lehr-/Unterrichtserfahrungen	Zeitstunden	0	1	4 %
		2	3	12 %
		4	2	8 %
		5	2	8 %
		6	3	12 %
		8	2	8 %
		10	1	4 %
		12	1	4 %
		16	1	4 %
		20	3	12 %
		24	1	4 %
		25	1	4 %
		30	1	4 %
		32	1	4 %
		50	1	4 %
		80	1	4 %
	240	1	4 %	

Tabelle A. 8: Beschreibung der Stichprobe: Ort und Fachbezug der Praktika der Studierenden gem. LABG 2009 (Bachelor of Arts)

Variable	Beschreibung	Wert	<i>N</i>	Prozent
Stichprobenumfang			26	
Studienfach des Berufsfeldpraktikums	Sozialwissenschaften		6	23 %
	Anderes		19	73 %
		fehlend	1	4 %
Praktikumsort Berufsfeldpraktikum	Außerschulische bildungsorientierte Einrichtung		5	19 %
	Schule		19	73 %
		fehlend	2	8 %

Tabelle A. 9: Beschreibung der Stichprobe: Absolvierte Praktika und Lehr-/Unterrichtserfahrungen der Studierenden gem. LABG 2002 (Erste Staatsprüfung)

Variable	Beschreibung	Wert	<i>N</i>	Prozent
Stichprobenumfang			24	
Praktika	Orientierungspraktikum (P_4)		6	25 %
	Fachpraktikum (P_5)		0	0 %
	Alle Praktika (P_4 und P_5)		18	75 %
Lehr-/Unterrichtserfahrungen	Zeitstunden	0	3	13 %
		2	1	4 %
		3	1	4 %
		5	1	4 %
		8	2	8 %
		10	2	8 %
		14	2	8 %
		15	2	8 %
		20	3	13 %
		50	4	17 %
		120	1	4 %
		200	1	4 %
				4 %

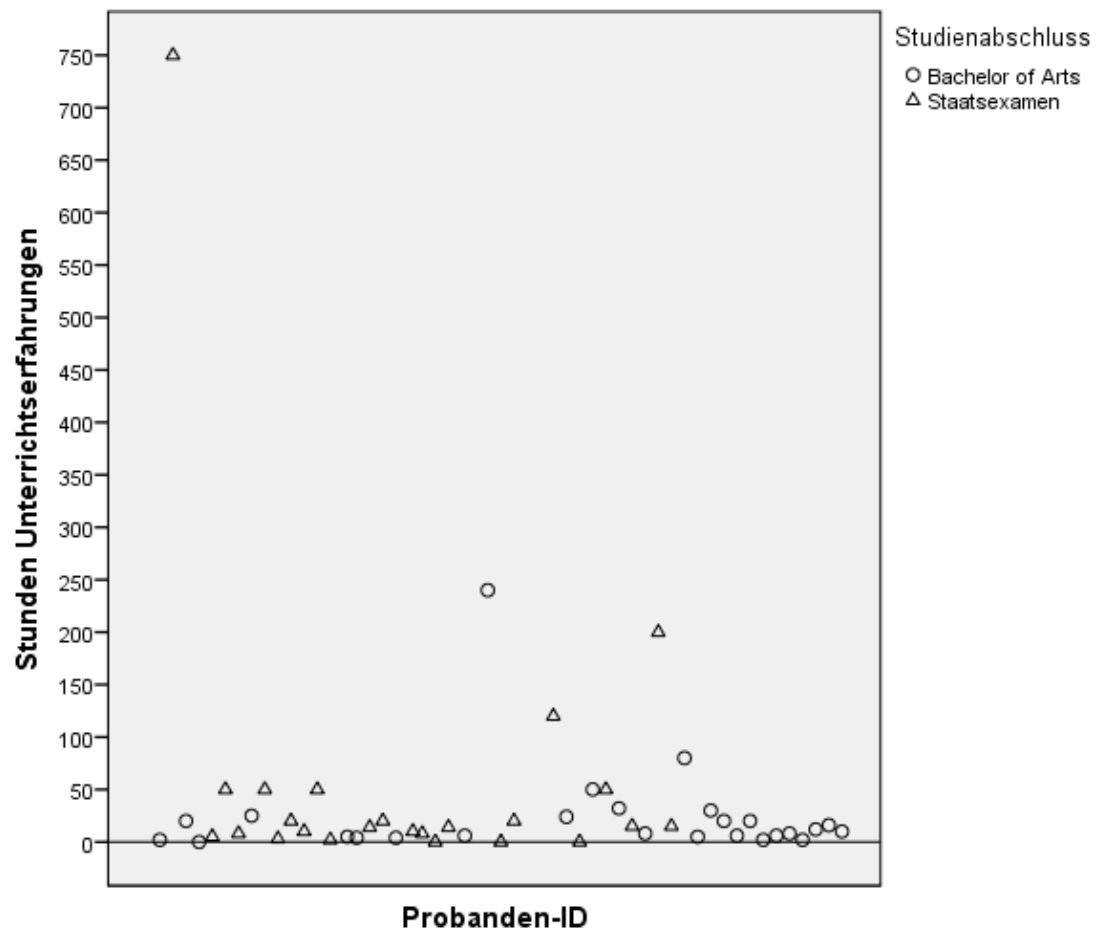


Abbildung A. 1: Beschreibung der Stichprobe: Streudiagramm (gruppiert) zur Verteilung der Angaben zu den Lehr-/Unterrichtserfahrungen der Probanden in Zeitstunden

Tabelle A. 10: Übersicht über technisch bedingt fehlende Werte im Computerprogramm vKid in der Skala zum deklarativen politikdidaktischen Wissen (FdW)

Item	fehlende Werte		Prozent
	AW < 2 Sek.	technisches Problem	
Stichprobenumfang	50	50	
FdW_1	0	0	0 %
FdW_2	0	0	0 %
FdW_3	0	0	0 %
FdW_4	0	0	0 %
FdW_5	0	0	0 %
FdW_6	1	0	2 %
FdW_7	0	0	0 %
FdW_8	0	0	0 %
FdW_9	0	0	0 %
FdW_10	1	0	2 %
FdW_11	0	0	0 %
FdW_12	0	0	0 %
FdW_13	0	0	0 %
FdW_14	1	0	2 %
FdW_15	0	0	0 %
FdW_16	0	0	0 %
FdW_17	0	0	0 %
FdW_18	0	0	0 %
FdW_19	0	0	0 %
FdW_20	1	0	2 %

Tabelle A. 11: Übersicht über technisch bedingt fehlende Werte im Computerprogramm vKid in der Skala zum prozeduralen politikdidaktischen Wissen (FpW)

Item	fehlende Werte		Prozent
	AW < 2 Sek.	technisches Problem	
Stichprobenumfang	50	50	
FpW-I_1	0	1	2 %
FpW-I_2	0	2	4 %
FpW-I_3	0	2	4 %
FpW-E_4	0	2	4 %
FpW-I_5	0	2	4 %
FpW-E_6	0	0	0 %
FpW-E_7	0	0	0 %
FpW-E_8	0	2	4 %
FpW-E_9	0	1	2 %
FpW-I_10	0	2	4 %

Tabelle A. 12: Übereinstimmungsmaße zwischen den Ratern in der Skala zum prozeduralen politikdidaktischen Wissen (FpW)

Item	N	κ (Cohen)	κ_w (Cohen)			ρ (Spearman)
			lineare Gewichtung	Ge- wichtung	quadratische Gewichtung	
Stichprobenumfang	50					
FpW-I_1	49	.72		.73	.74	.73**
FpW-I_2	48	.72		.69	.65	.68**
FpW-I_3	48	.79		.79	.80	.83**
FpW-E_4	48	.65		.70	.77	.73**
FpW-I_5	48	.66		.67	.69	.67**
FpW-E_6	50	.62		.66	.71	.76**
FpW-E_7	50	.63		.66	.69	.70**
FpW-E_8	48	.70		.71	.73	.74**
FpW-E_9	49	.71		.71	.71	.72**
FpW-I_10	48	.66		.66	.66	.70**

** Die Korrelation ist auf dem .01 Niveau (zweiseitig) signifikant.

Anmerkung: Als relative Distanz zwischen den Stufen der 0/1/2-Skala (Ordinalniveau) für die Berechnung des gewichteten Cohens Kappa (κ_w) wird der Wert *eins* festgelegt.

Tabelle A. 13: Itemkennwerte zur Skala zum Fachwissen (FW) mit Anzahl gültiger Fälle, Schwierigkeitsindex p_i , Itemtrennschärfe r_{it} sowie Verteilung der Testleistung (absolute Häufigkeiten, Mittelwert M und Standardabweichung SD)

Item	N	p_i	r_{it}	Häufigkeiten	
				1 (richtig)	0 (falsch)
Stichprobenumfang	50				
FW_1	50	.68	.43	34	16
FW_2	50	.56	.39	28	22
FW_3	50	.84	-.12	42	8
FW_4	50	.80	.51	40	10
FW_5	50	.76	.51	38	12
FW_6	50	.94	-.01	47	3
FW_7	50	.50	.19	25	25
FW_8	50	.60	.32	30	20
FW_9	50	.22	.40	11	39
FW_10	50	.92	.49	46	4
FW_11	50	1.00	.00	50	0
FW_12	50	.92	.23	46	4
FW_13	50	.40	.43	20	30
FW_14	50	.90	.41	45	5
FW_15	50	.96	.31	48	2
FW_16	50	.50	-.13	25	25
FW_17	50	.92	.02	46	4
FW_18	50	.80	.53	40	10
FW_19	50	.70	.38	35	15
FW_20	50	.30	.09	15	35
FW_21	50	.86	.29	43	7
FW_22	50	.38	-.08	19	31
FW_23	50	.26	.36	13	37
FW_24	50	.44	.54	22	28
FW_25	50	.96	.02	48	2
FW_26	50	.56	-.02	28	22

Fortsetzung

Item	N	p_i	r_{it}	Häufigkeiten	
				1 (richtig)	0 (falsch)
Stichprobenumfang	50				
FW_27	50	.18	-.27	9	41
FW_28	50	.52	.26	26	24
FW_29	50	.52	.05	26	24
FW_30	50	.42	.50	21	29
FW_31	50	.56	.40	28	22
FW_32	50	.54	.53	27	23
FW_33	50	.90	.16	45	5
FW_34	50	.08	.29	4	46
FW_35	50	.94	.14	47	3
FW_36	50	.40	-.01	20	30
FW_37	50	.72	.21	36	14
FW_38	50	.32	.33	16	34
FW_39	50	.60	.70	30	20
FW_40	50	.10	.42	5	45

Tabelle A. 14: Itemkennwerte zur Skala zum deklarativen politikdidaktischen Wissen (FdW) mit Anzahl gültiger Fälle, Schwierigkeitsindex p_i . Itemtrennschärfe r_{it} sowie Verteilung der Testleistung (absolute Häufigkeiten, Mittelwert M und Standardabweichung SD).

Item	N	p_i	r_{it}	Häufigkeiten	
				1 (richtig)	0 (falsch)
Stichprobenumfang	50				
FdW_1	50	.44	.09	22	28
FdW_2	50	.52	.46	26	24
FdW_3	50	.84	.06	42	8
FdW_4	50	.44	.25	22	28
FdW_5	50	.86	.28	43	7
FdW_6	49	.65	.22	32	17
FdW_7	50	.52	.21	26	24
FdW_8	50	.82	-.07	41	9
FdW_9	50	.38	.16	19	31
FdW_10	49	.88	.07	43	6
FdW_11	50	.52	.17	26	24
FdW_12	50	.50	-.14	25	25
FdW-L_13	50	.96	.43	48	2
FdW-L_14	49	.86	.26	42	7
FdW-L_15	50	.94	-.03	47	3
FdW-L_16	50	.24	-.07	12	38
FdW-L_17	50	.22	.09	11	39
FdW-L_18	50	.64	.18	32	18
FdW-L_19	50	.58	.42	29	21
FdW-L_20	49	.86	.28	42	7

Tabelle A. 15: Itemkennwerte zur Skala zum prozeduralen politikdidaktischen Wissen mit Anzahl gültiger Fälle, Schwierigkeitsindex p_i , Itemtrennschärfe r_{it} sowie Verteilung der Testleistung (absolute Häufigkeiten, Mittelwert M und Standardabweichung SD)

Item	N	p_i	r_{it}	Häufigkeiten		
				2	1	0
Stichprobenumfang	50					
FpW-I_1	49	.36	.12	1	33	15
FpW-I_2	48	.43	.28	4	33	11
FpW-I_3	48	.23	.22	1	20	27
FpW-E_4	48	.40	.20	7	24	17
FpW-I_5	48	.31	.23	1	28	19
FpW-E_6	50	.38	.38	3	32	15
FpW-E_7	50	.34	.39	2	30	18
FpW-E_8	48	.33	.33	2	28	18
FpW-E_9	49	.26	.36	0	25	24
FpW-I_10	48	.39	.08	0	37	11

Anmerkung: Die Werte bei der Angabe der absoluten Häufigkeiten repräsentieren die Skalenwerte 2 = *adäquat*, 1 = *teilweise adäquat* und 0 = *inadäquat*.

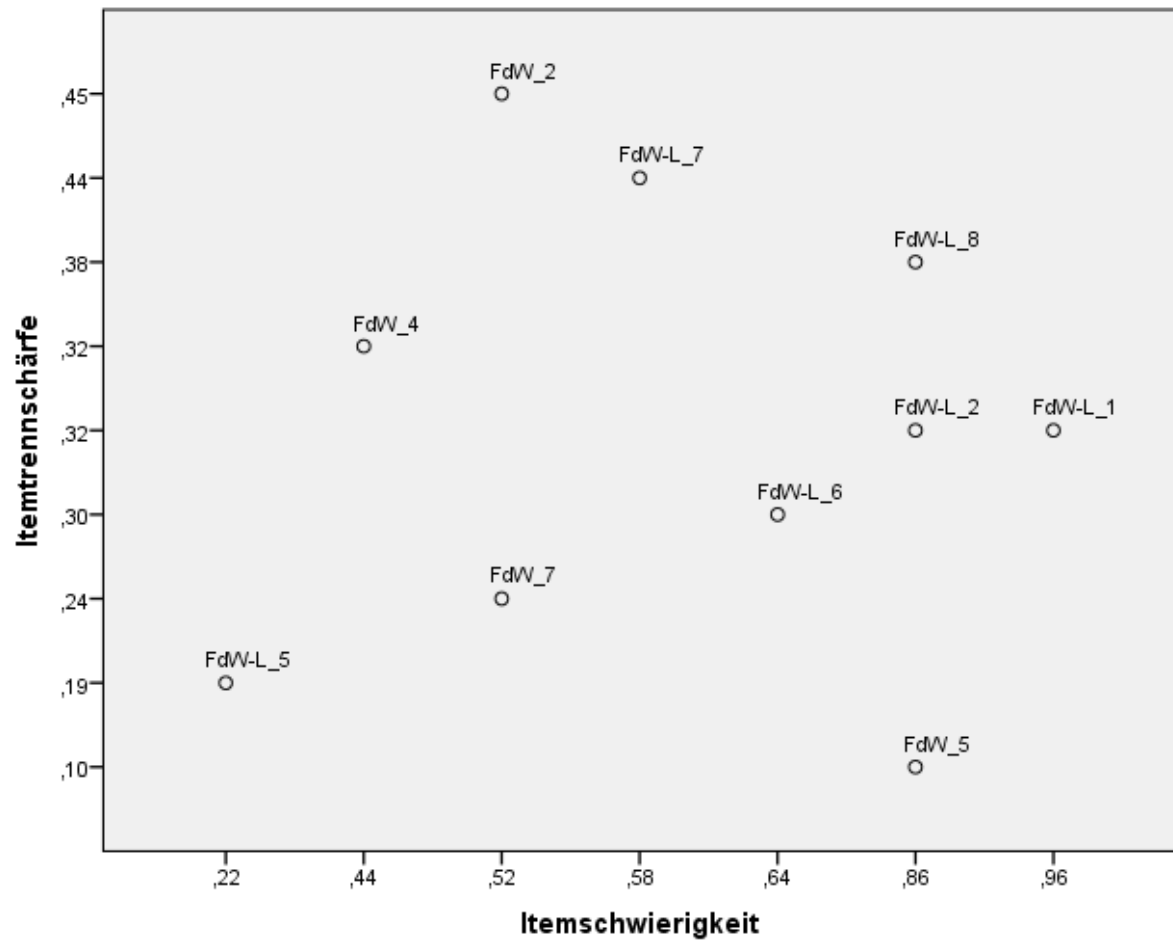


Abbildung A. 2: Zusammenhang zwischen Itemschwierigkeit und Trennschärfe in den vorliegenden Daten zur Skala zum deklarativen politikdidaktischen Wissen unter Berücksichtigung ausgeschlossener Items (FdW_rev)

Anmerkung: Die Itemtrennschärfe weist die korrigierte Item-Skala-Korrelation unter Berücksichtigung der ausgeschlossenen Items aus; es wurden 10 Items ausgeschlossen.

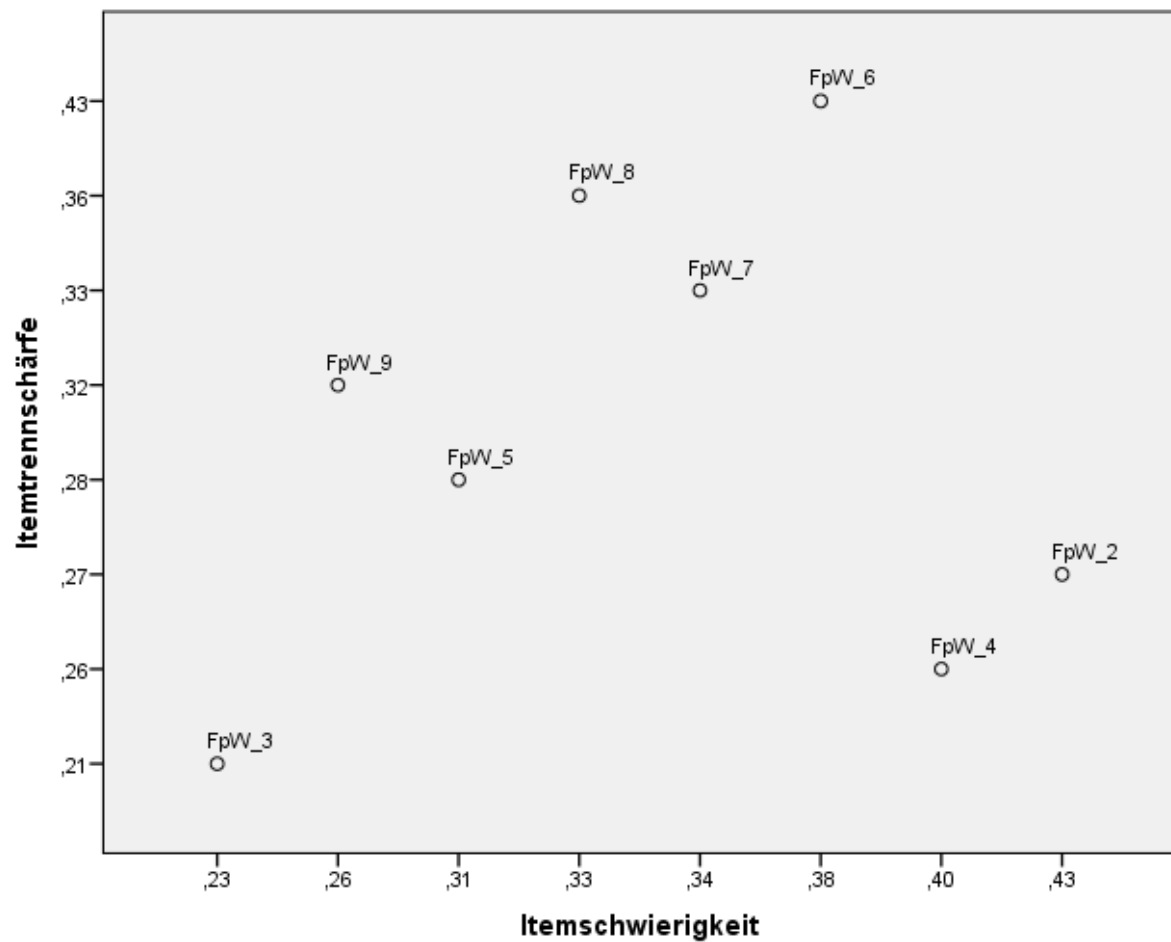


Abbildung A. 3: Zusammenhang zwischen Itemschwierigkeit und Trennschärfe in den vorliegenden Daten zur Skala zum prozeduralen politikdidaktischen Wissen unter Berücksichtigung ausgeschlossener Items (FpW_rev)

Anmerkung: Die Itemtrennschärfe weist die korrigierte Item-Skala-Korrelation unter Berücksichtigung der ausgeschlossenen Items aus; es wurden zwei Items ausgeschlossen.

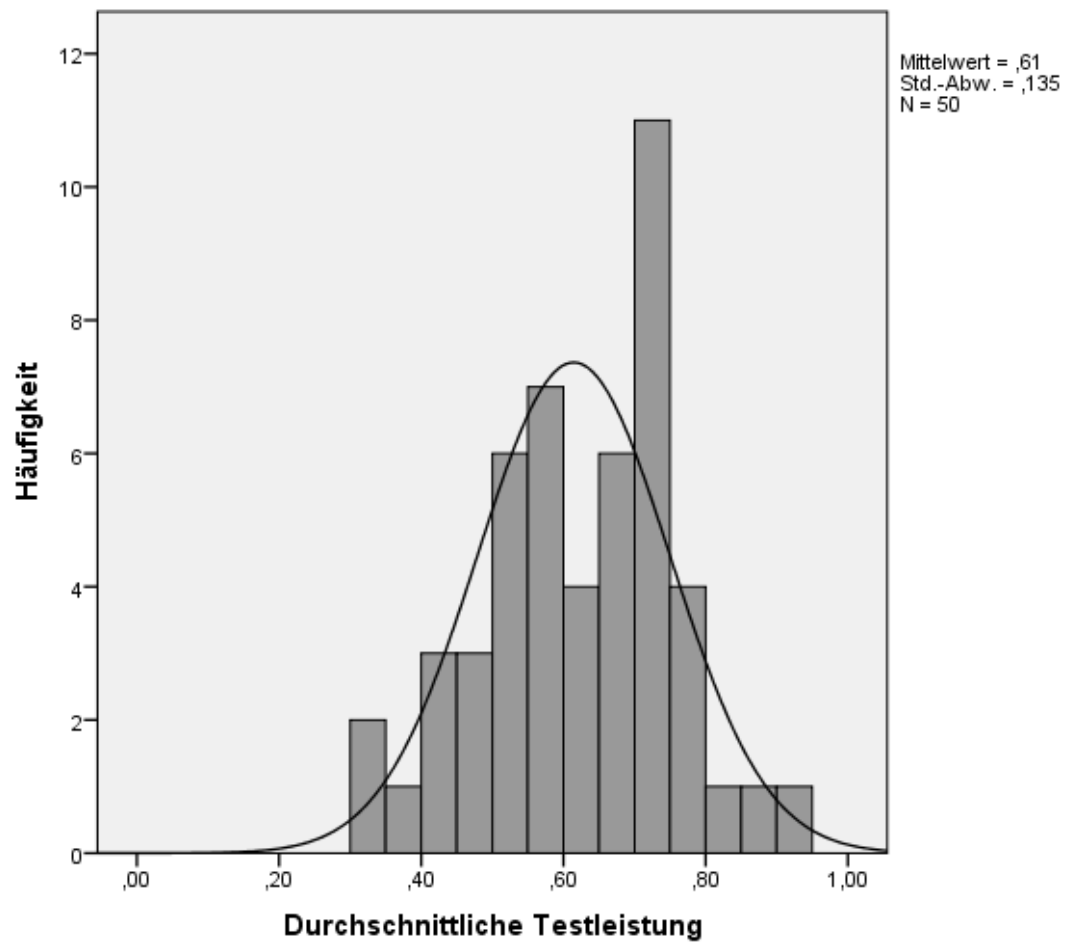


Abbildung A. 4: Verteilung der durchschnittlichen Testleistungen der Probanden (Personenfähigkeit) in der Skala zum Fachwissen (FW) mit Normalverteilungskurve

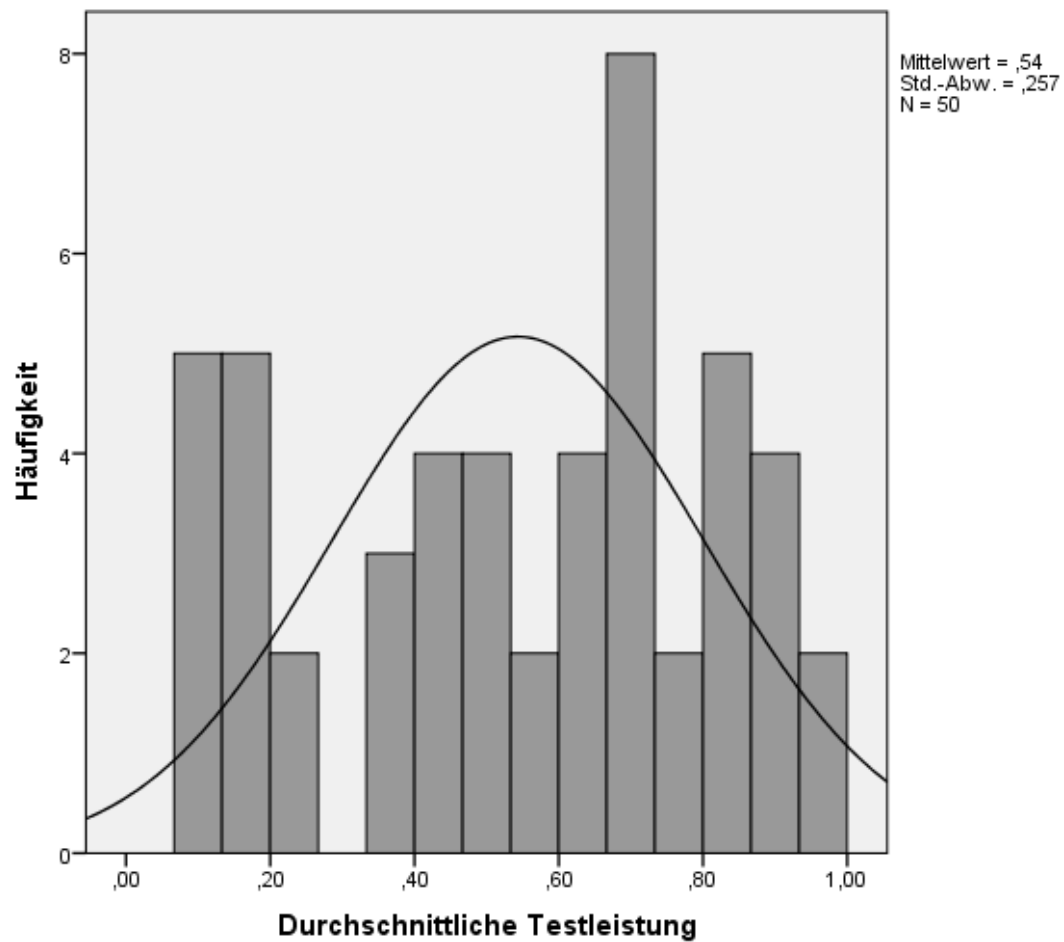


Abbildung A. 5: Verteilung der durchschnittlichen Testleistungen der Probanden (Personenfähigkeit) in der revidierten Skala zum Fachwissen (FW_rev) mit Normalverteilungskurve

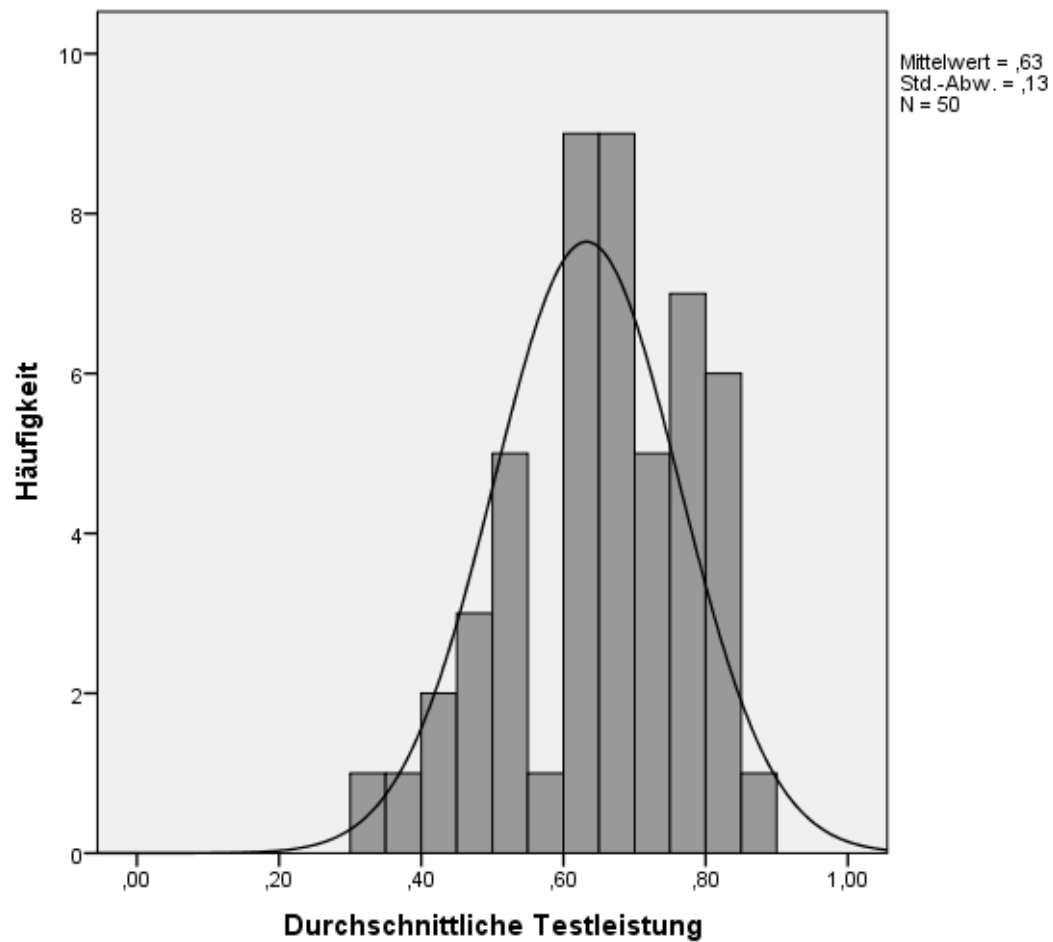


Abbildung A. 6: Verteilung der durchschnittlichen Testleistungen der Probanden (Personenfähigkeit) in der Skala zum deklarativen politikdidaktischen Wissen (FdW) mit Normalverteilungskurve

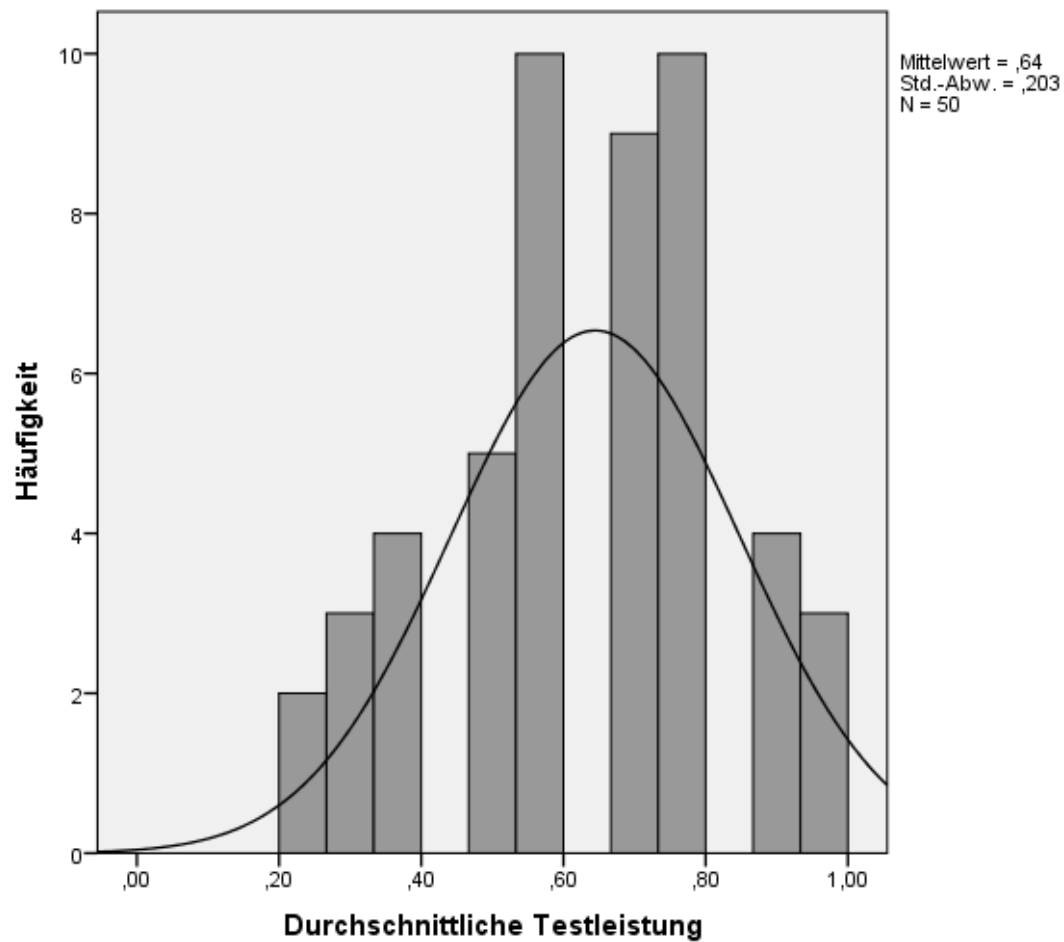


Abbildung A. 7: Verteilung der durchschnittlichen Testleistungen der Probanden (Personenfähigkeit) in der revidierten Skala zum deklarativen politikdidaktischen Wissen (FdW_rev) mit Normalverteilungskurve

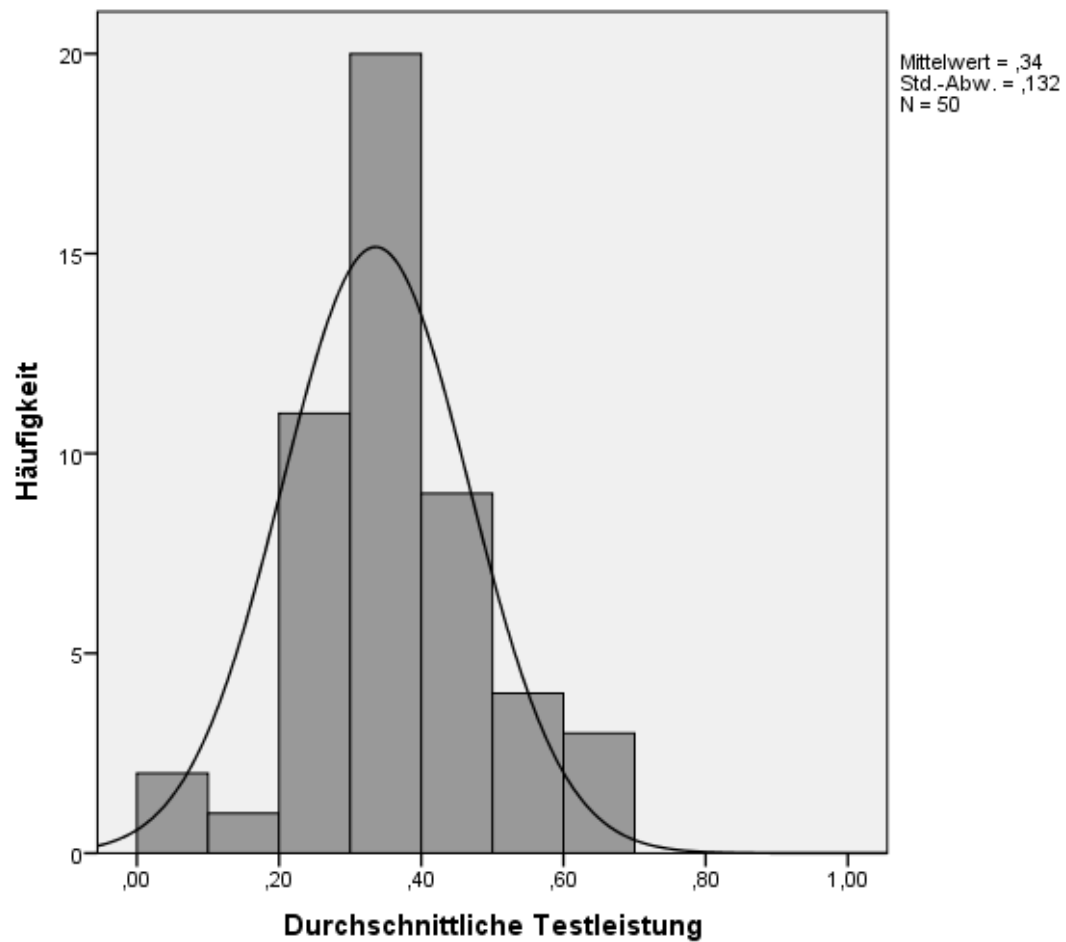


Abbildung A. 8: Verteilung der durchschnittlichen Testleistungen der Probanden (Personenfähigkeit) in der Skala zum prozeduralen politikdidaktischen Wissen (FpW) mit Normalverteilungskurve

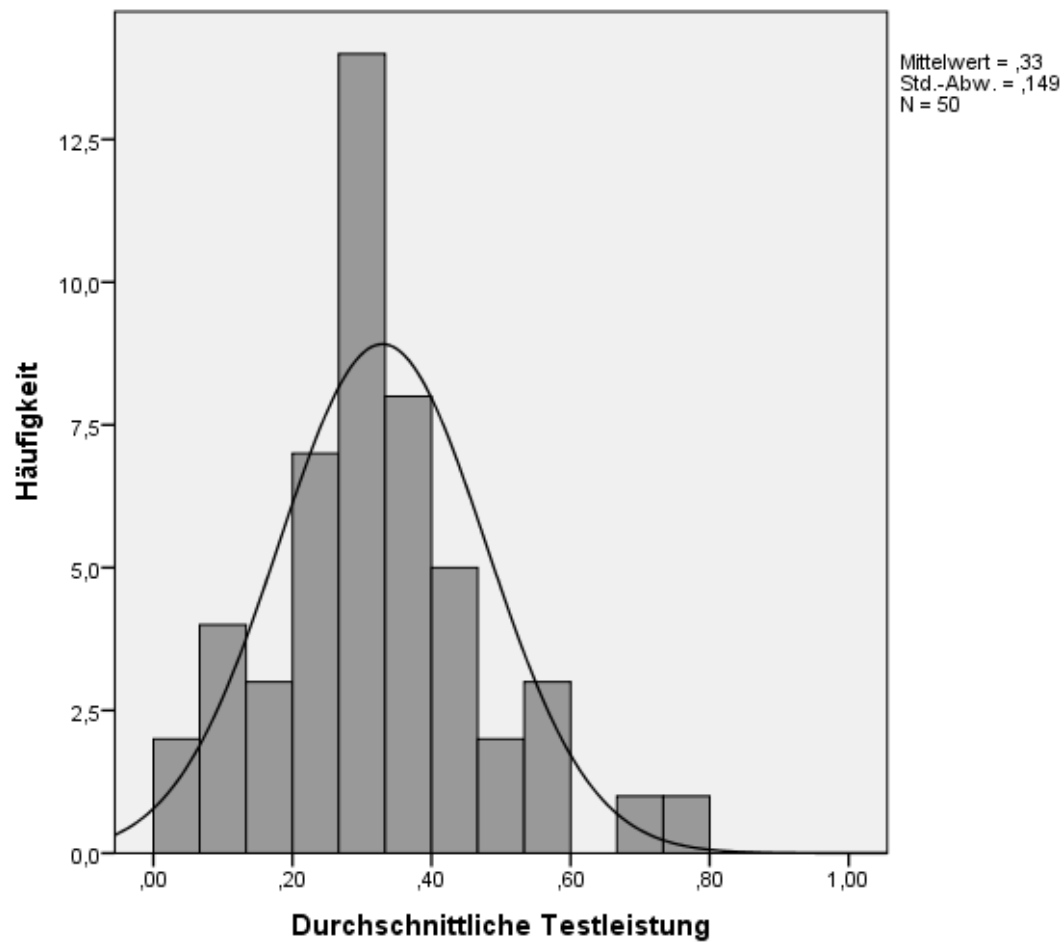


Abbildung A. 9: Verteilung der durchschnittlichen Testleistungen der Probanden (Personen-
fähigkeit) in der revidierten Skala zum prozeduralen politikdidaktischen Wissen (FpW_rev)
mit Normalverteilungskurve

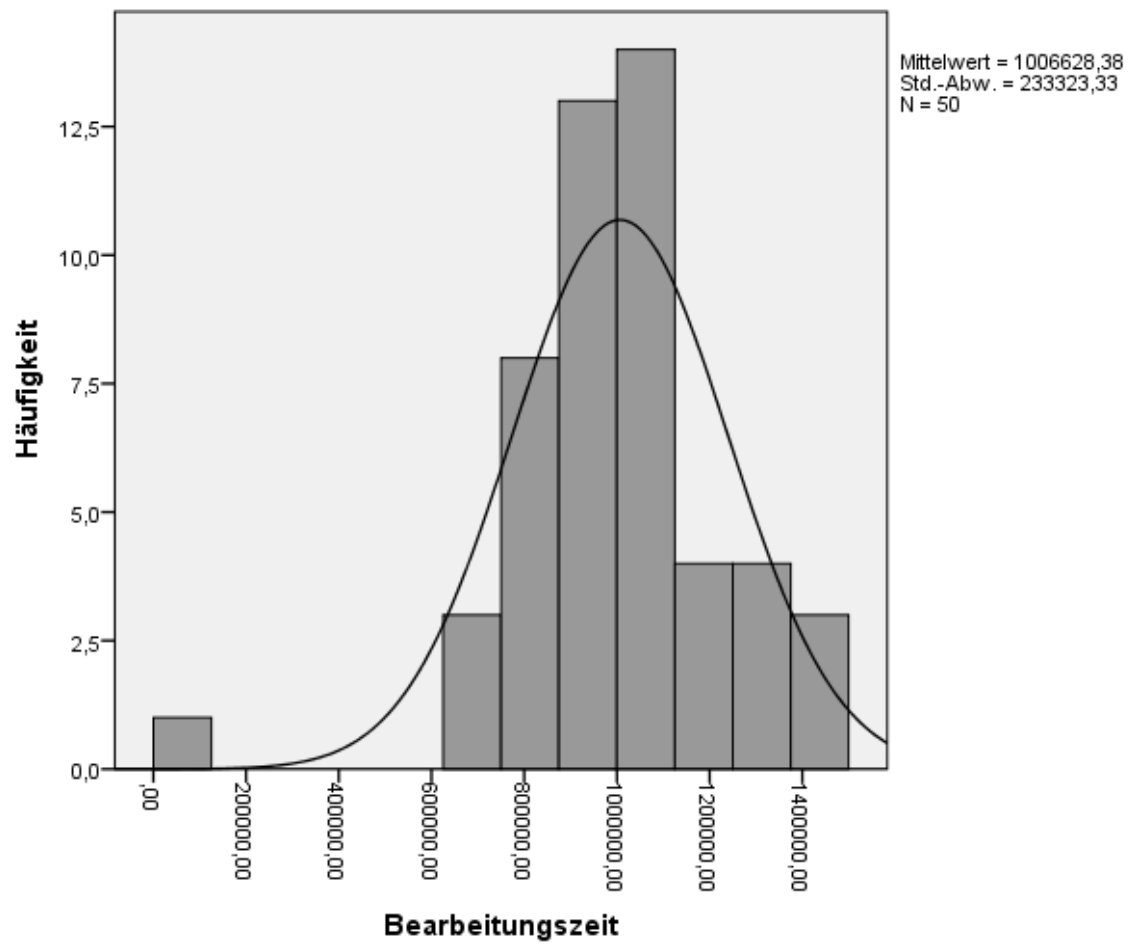


Abbildung A. 10: Histogramm mit Normalverteilungskurve zu den kumulierten Bearbeitungszeiten (in Millisekunden) der Probanden in der Skala zum prozeduralen politikdidaktischen Wissen (FpW)

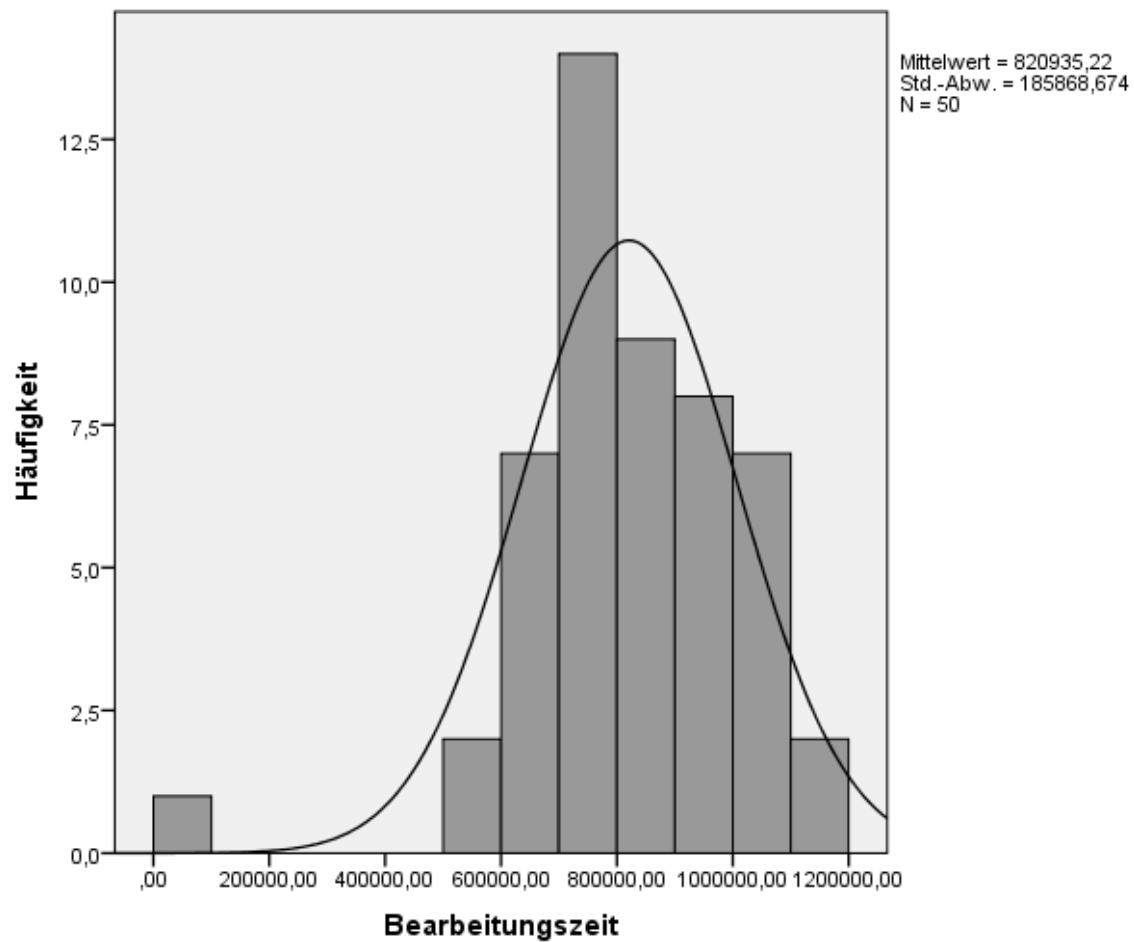


Abbildung A. 11: Histogramm mit Normalverteilungskurve zu den kumulierten Bearbeitungszeiten (in Millisekunden) der Probanden in der revidierten Skala zum prozeduralen politikdidaktischen Wissen (FpW_rev)